PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/48, C07K 14/15, A61K 39/21, C07K 16/10, C12Q 1/70, A01K 67/027

A1

(11) International Publication Number: WO 97/40167

(43) International Publication Date: 30 October 1997 (30.10.97)

(21) International Application Number: PCT/GB97/01087

(22) International Filing Date: 18 April 1997 (18.04.97)

(30) Priority Data:

9608164.1	19 April 1996 (19.04.96)	GB
9702668.6	10 February 1997 (10.02.97)	GB

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(71) Applicants *(for all designated States except US)*: Q-ONE BIOTECH LTD. [GB/GB]; Todd Campus, West of Scotland Science Park, Glasgow G20 0XA (GB). IMUTRAN LTD. [GB/GB]; Douglas House, 18 Trumpington Road, Cambridge CB2 2AH (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: GALBRAITH, Daniel, Norman [GB/GB]; 26 Don Street, Glasgow G33 2DH (GB). HAWORTH, Christine [GB/GB]; 139 Marlborough Avenue, Broomhill, Glasgow G11 7JE (GB). LEES, Gillian, Margaret [GB/GB]; 5 Kirkvale Drive, Newton Mearns, Glasgow G77 5HD (GB). SMITH, Kenneth, Thomas [GB/GB]; 5 Kirkvale Drive, Newton Mearns, Glasgow G77 5HD (GB).

(74) Agents: McCALLUM, William, Potter et al.; Cruikshank & Fairweather, 19 Royal Exchange Square, Glasgow G1 3AE (GB).

(54) Title: PORCINE RETROVIRUS

(57) Abstract

The present invention provides porcine retrovirus (PoEV) polynucleotide fragments, particularly those encoding at least one PoEV expression product, a recombinant vector comprising such a polynucleotide fragment or fragments, use of PoEV polynucleotide fragments in the detection of native PoEV, a host cell containing at least one PoEV polynucleotide fragment or recombinant vector, PoEV polypeptides, antibodies immuno-reactive with PoEV polypeptides, pharmaceutical compositions comprising recombinant PoEV polypeptides for use as prophylactic and/or therapeutic agents and uses of PoEV polynucleotide fragments and/or polypeptides in medicine, including veterinary medicine and in the preparation of medicaments for use in medicine.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Porcine Retrovirus

The present invention relates *inter alia* to porcine retrovirus (PoEV) fragments, in particular polynucleotide fragments encoding at least one porcine retrovirus expression product, a recombinant vector comprising at least one polynucleotide fragment, use of PoEV polynucleotide fragments in the detection of native porcine retrovirus, a host cell containing at least one PoEV polynucleotide fragment or a recombinant vector comprising at least one PoEV polynucleotide fragment, PoEV polypeptides, antibodies immuno-reactive with PoEV polypeptides, pharmaceutical compositions comprising recombinant PoEV polypeptides for use as prophylactic and/or therapeutic agents and uses of PoEV polynucleotide fragments and/or polypeptides in medicine, including veterinary medicine and in the preparation of medicaments for use in medicine, including veterinary medicine.

Porcine retrovirus (PoEV) is an endogenous (genetically acquired) retrovirus isolated from pigs and expressed in cell lines derived from porcine material. There are no known pathogenic effects associated with the virus per se in its natural host although the virus appears to be associated with lymphomas in pigs and related viruses are associated with leukaemias and lymphomas in other species. The virus has been reported to infect cells from a variety of non-porcine origins and is, therefore, designated as a xenotropic, amphotropic or polytrophic virus (Lieber MM, Sherr CJ. Benveniste RE and Todaro

GJ. 1975; Strandstrom H, Verjalainen P, Moening V, Hunsmann G, Schwarz H, and Schafer W. 1974; Todaro GJ, Benveniste RE, Lieber MM and Sherr CJ. 1974). The observation that the above viruses may have the potential to infect humans and have a pathogenic effect suggests that the issue of porcine retroviruses must be addressed in the context of xenotransplanting pig tissues or cells. Therefore, information on the properties of PoEV and the development of diagnostic reagents, molecular engineering tools and potential vaccine materials is of paramount importance for example in xenotransplantation technology and the like.

It is an object of the present invention to obviate and/or mitigate against at least some of the above disadvantages.

In one aspect the present invention provides an isolated PoEV polynucleotide fragment:

(a) encoding at least one porcine retrovirus (PoEV) expression product;

(b) encoding a physiologically active and/or immunogenic derivative of said expression product; or

(c) which is complementary to a polynucleotide sequence as defined in (a) or (b).

Preferably, the polynucleotide fragment encodes the gag gene (gag), polymerase gene (*pol*) and/or envelope (*env*) gene of PoEV. Thus, said expression product can be the virion core polypeptides (GAG) and polymerase (POL) and/or envelope (ENV) polypeptides of PoEV. Thus, the invention further provides a recombinant PoEV virion core, polymerase and/or envelope polypeptide.

"Polynucleotide fragment" as used herein refers to a chain of nucleotides such as deoxyribose nucleic acid (DNA) and

transcription products thereof, such as RNA. Naturally, the skilled addressee will appreciate the whole naturally occurring PoEV genome is not included in the definition of polynucleotide fragment.

The polynucleotide fragment can be isolated in the sense that it is substantially free of biological material with which the whole genome is normally associated *in vivo*. The isolated polynucleotide fragment may be cloned to provide a recombinant molecule comprising the polynucleotide fragment. Thus, "polynucleotide fragment" includes double and single stranded DNA, RNA and polynucleotide sequences derived therefrom, for example, subsequences of said fragment and which are of any desirable length. Where a nucleic acid is single stranded then both a given strand and a sequence complementary thereto is within the scope of the present invention.

In general, the term "expression product" refers to both transcription and translation products of said polynucleotide fragments. When the expression product is a "polypeptide" (i.e. a chain or sequence of amino acids displaying a biological and/or immunological activity substantially similar to the biological and/or immunological activity of PoEV virion core, polymerase and/or envelope protein), it does not refer to a specific length of the product as such. Thus, the skilled addressee will appreciate that "polypeptide" encompasses *inter alia* peptides, polypeptides and proteins of PoEV. The polypeptide if required, can be modified *in vivo* and *in vitro*, for example by glycosylation, amidation, carboxylation, phosphorylation and/or post-translational cleavage.

Polynucleotide fragments comprising portions encompassing the PoEV genome, and derived from retrovirus particles released from a reverse transcriptase-positive porcine kidney cell line PK-15, have been molecularly cloned into a plasmid vector. This was achieved by synthesising cDNAs of PoEV RNA genomes which were recovered from porcine kidney cells expressing the endogenous virus. The cDNA was cloned into a plasmid vector and the isolated PoEV DNA fragment determined (see Figures 1,2 and 3). The sequence of the sequence identified in Figure 1 was the earliest determined sequence, followed by the sequence in Figure 2 and lastly by the most recently revised sequence shown in Figure 3. An additional study has been carried out to determine whether or not the human cell line "Raji" was susceptible to infection with the PoEV present in porcine kidney cells (PK15). A raji clone has now been obtained and the DNA sequence of its *env* gene region has been determined (see Figure 4).

The DNA fragment of Figure 3 was shown to encode three open reading frames (ORFs) of 524, 1194 and 656 amino acids respectively.

A comparison of the amino acid sequence against previously sequenced retroviruses from other species indicated that novel retrovirus cDNA had been cloned. Sequence analysis using the Lasergene software from DNASTAR Inc. showed that homologies were observed between the cloned PoEV DNA and the majority of retroviruses and that the closest homologies were to gibbon leukaemia virus (GaLV) in the polymerase (*pol*) and (*env*) regions of the pro-virus.

The first open reading frame ORF of Figure 3 (nucleotides 588-

2162) is predicted to encode the PoEV virion core polypeptide (*gag* gene). The second ORF (nucleotides 2163-5747) is predicted to encode the PoEV polymerase polypeptide (*pol* gene). The third ORF (nucleotides 5620-7590) is predicted to encode the PoEV envelope polypeptide (*env* gene). The skilled addressee will appreciate that it is possible to genetically manipulate the polynucleotide fragment or derivatives thereof, for example to clone the gene by recombinant DNA techniques generally known in the art and to express the polypeptides encoded thereby in vitro and/or in vivo. DNA fragments having the polynucleotide sequence depicted in Figures 1,2,3 and/or 4 or DNA/RNA derivatives thereof, may be used as a diagnostic tool or as a reagent for detecting PoEV nucleic acid in nucleic acids from donor animals or as a vaccine.

Preferred fragments of this aspect of the invention are polynucleotide fragments encoding: (a) at least one of the one to three polypeptides having an amino acid sequence which is shown in Figures 1,2,3 and/or 4 (b) encoding a polypeptide which is a physiologically active and/or immunogenic derivative of at least one of the polypeptides defined in (a); or (c) which is complementary to a polynucleotide sequence as defined above; or polynucleotide fragments: (a) comprising at least one of the ORFs shown in Figures 1,2,3 and/or 4 or comprising a corresponding RNA sequence; (b) comprising a sequence having substantial nucleotide sequence identity with a sequence as described in (a) above; or (c) comprising a sequence which is complementary to a sequence as described in (a) or (b) above. It is to be understood that the term "substantial sequence identity" is taken to mean at

least 50% (preferably at least 75%, at least 90%, or at least 95%) sequence identity.

The polynucleotide fragment of the present invention may be used to examine the expression and/or presence of the PoEV virus in donor animals and cells, tissues or organs derived from the donor animals to see if they are suitable for xenotransplantation (i.e. PoEV free). In addition, the recipients of pig cells, tissues or organs can be examined for the presence and/or expression of PoEV virus directly or by co-culture or infection of susceptible detector cells.

A polynucleotide fragment of the present invention may be used to identify polynucleotide sequences within the PoEV genome which are PoEV specific (i.e. it is not necessary for the complete PoEV genome to be identified). Such PoEV specific polynucleotide sequences may be used to identify PoEV nucleic acid in samples, such as transplanted cells, tissues or organs and may be included in a definitive test for PoEV.

Thus, the present invention further provides an isolated PoEV polynucleotide fragment capable of specifically hybridising to a PoEV polynucleotide sequence. In this manner, the present invention provides probes and/or primers for use in *ex vivo* and/or *in situ* PoEV virus detection and expression studies. Typical detection studies include polymerase chain reaction (PCR) studies, hybridisation studies, or sequencing studies. In principle any PoEV specific polynucleotide sequence from the above identified PoEV sequence may be used in detection and/or expression studies.

"Capable of specifically hybridising" is taken to mean that

said polynucleotide fragment preferably hybridises to a PoEV polynucleotide sequence in preference to polynucleotide sequences of other virus, animal (especially porcine or human sequences) and/or other species. In a preferment the PoEV fragment specifically binds to a native PoEV polynucleotide sequence or a part thereof.

The invention includes polynucleotide sequence(s) which are capable of specifically hybridising to a PoEV polynucleotide sequence or to a part thereof without necessarily being completely complementary to said PoEV polynucleotide sequence or fragment thereof. For example, there may be at least 50% preferably at least 75%, most preferably at least 90% or at least 95% complementarity. Of course, in some cases the sequences may be exactly complementary (100% complementary) or nearly so (e.g. there may be less than 10, preferably less than 5 mismatches). Thus, the present invention also provides anti-sense or complementary nucleotide sequence(s) which is/are capable of specifically hybridising to the disclosed DNA sequence. If a PoEV specific polynucleotide is to be used as a primer in PCR and/or sequencing studies, the polynucleotide must be capable of hybridising to PoEV nucleic acid and capable of initiating chain extension from 3' end of the polynucleotide, but not able to correctly initiate chain extension from non PoEV sequences (especially from human, or non-PoEV porcine sequences).

If a PoEV specific test polynucleotide sequence is to be used in hybridisation studies, to test for the presence of PoEV nucleic acid in a sample, the test polynucleotide should preferably remain hybridised to a sample polynucleotide under

stringent conditions. If desired, either the test or sample polynucleotide may be immobilised. Generally the test polynucleotide sequence is at least 10 or at least 50 bases in length. It may be labelled by suitable techniques known in the art. Preferably the test polynucleotide sequence is at least 200 bases in length and may even be several kilobases in length. Thus, either a denatured sample or test sequence can be first bound to a support. Hybridization can be effected at a temperature of between 50 and 70°C in double strength SSC (2xNaCl 17.5g/l and sodium citrate (SC) at 8.8g/l) buffered saline containing 0.1% sodium dodecyl sulphate (SDS). This can be followed by rinsing of the support at the same temperature but with a buffer having a reduced SSC concentration. Depending upon the degree of stringency required, and thus the degree of similarity of the sequences, such reduced concentration buffers are typically single strength SSC containing 0.1%SDS, half strength SSC containing 0.1%SDS and one tenth strength SSC containing 0.1%SDS. Sequences having the greatest degree of similarity are those the hybridisation of which is least affected by washing in buffers of reduced concentration. It is most preferred that the sample and inventive sequences are so similar that the hybridisation between them is substantially unaffected by washing or incubation in one tenth strength sodium citrate buffer containing 0.1%SDS.

PoEV specific oligonucleotides may be designed to specifically hybridise to PoEV nucleic acid. They may be synthesised, by known techniques and used as primers in PCR or sequencing reactions or as probes in hybridisations designed to

detect the presence of PoEV material in a sample. The oligonucleotides may be labelled by suitable labels known in the art, such as, radioactive labels, chemiluminescent labels or fluorescent labels and the like. Thus, the present invention also provides PoEV specific oligonucleotide probes and primers.

The term "oligonucleotide" is not meant to indicate any particular length of sequence and encompasses nucleotides of preferably at least 10b (e.g. 10b to 1kb) in length, more preferably 12b-500b in length and most preferably 15b to 100b.

The PoEV specific oligonucleotides may be determined from the PoEV sequences shown in Figure 1 and may be manufactured according to known techniques. They may have substantial sequence identity (e.g. at least 50%, at least 75%, at least 90% or at least 95% sequence identity) with one of the strands shown therein or an RNA equivalent, or with a part of such a strand. Preferably such a part is at least 10, at least 30, at least 50 or at least 200 bases long. It may be an ORF or a part thereof.

Oligonucleotides which are generally greater than 30 bases in length should preferably remain hybridised to a sample polynucleotide under one or more of the stringent conditions mentioned above. Oligonucleotides which are generally less than 30 bases in length should also preferably remain hybridised to a sample polynucleotide but under different conditions of high stringency. Typically the melting temperature of an oligonucleotide less than 30 bases may be calculated according to the formula of; 2°C for every A or T, plus 4°C for every G or C, minus 5°C. Hybridisation may take place at or around the calculated melting temperature for any particular

oligonucleotide, in 6 x SSC and 1% SDS. Non specifically hybridised oligonucleotides may then be removed by stringent washing, for example in 3 x SSC and 0.1% SDS at the same temperature. Only substantially similar matched sequences remain hybridised i.e. said oligonucleotide and corresponding PoEV nucleic acid.

When oligonucleotides of generally less than 30 bases in length are used in sequencing and/or PCR studies, the melting temperature may be calculated in the same manner as described above. The oligonucleotide may then be allowed to anneal or hybridise at a temperature around the oligonucleotides calculated melting temperature. In the case of PCR studies the annealing temperature should be around the lower of the calculated melting temperatures for the two priming oligonucleotides. It is to be appreciated that the conditions and melting temperature calculations are provided by way of example only and are not intended to be limiting. It is possible through the experience of the experimenter to vary the conditions of hybridisation and thus anneal/hybridise oligonucleotides at temperatures above their calculated melting temperature. Indeed this can be desirable in preventing so-called non-specific hybridisation from occurring.

It is possible when conducting PCR studies to predict an expected size or sizes of PCR product(s) obtainable using an appropriate combination of two or more PoEV oligonucleotides, based on where they would hybridise to the sequence in Figure 1. If, on conducting such a PCR on a sample of PoEV DNA, a fragment of the predicted size is obtained, then this is predictive that

the DNA is PoEV.

The present invention also encompasses PoEV detection kits including at least one oligonucleotide which is PoEV specific, as well as any necessary reaction reagents, washing reagents, detection reagents, signal producing agents and the like for use in the test formats outlined above.

In a further aspect there is also provided use of a PoEV specific polynucleotide in the detection of PoEV in a sample.

In a yet further aspect there is provided use of a PoEV specific polynucleotide in a PCR for the detection of PoEV in a sample.

The skilled addressee will appreciate how polynucleotide fragments may be designed and used as primers/probes in polymerase chain reaction (PCR) experiments or Southern analysis (i.e. hybridisation studies) for detecting the presence or otherwise of PoEV polynucleotide in the nucleic acid of pigs or in cell, tissue or organ samples taken from pigs (e.g. from potential transplant organs such as liver, kidney and heart). Such cells, tissues or organs can be derived from transgenic animals produced as described in EP-A-0493852, or by other means known in the art. Thus the cells, tissues or organs of transgenic pigs can be associated with one or more homologous complement restriction factors active in humans to prevent/reduce activation of complement.

Furthermore the polynucleotide fragments of the present invention can be used to analyze the genetic organisation of endogenous PoEV located in the animal cell genome in pigs thus permitting the screening of herds of pigs for altered provirus

and genomic loci (e.g. non-expressed provirus loci). Such a screening method would facilitate, for example, screening in a population of animals which are bred to lack expressed provirus and genomic loci and/or loci that do not encode infectious virus particles.

Reagents may also be developed from said polynucleotide fragments as aids to develop pigs that do not express an infectious, PoEV capable of infecting humans. Such pigs could still contain partial defective genomes that could result in the expression of non-infectious particles, viral proteins or viral mRNA. Alternatively, it may be possible to use constructs derived from the PoEV polynucleotide sequence to act as insertional mutagens to knockout the productive infectious PoEV in embryos, embryonic stem cells, or cells containing totipotential nuclei capable of forming a viable embryo. Thus *gag*, *pol* and/or *env* gene "knockouts" may be constructed to allow development of breeding programmes in pigs whereby endogenous PoEV is substantially prevented or reduced. For example the nucleotide sequence of PoEV can be manipulated e.g. by deletion of a coding sequence in vitro and the resulting construct used to replace the natural PoEV sequence by recombination. Thus, the proviral genome can be rendered inactive in the porcine cells. The knockouts can be manipulated into embryos and/or stem cells and if required manipulated nuclei can be transferred from target cells to germ cells using micromanipulation techniques well known in the art. The invention also extends to animals derived from such germ cells.

Thus, transgenic pigs may be produced containing anti-sense

constructs and/or ribozyme constructs capable of downregulating the expression of viral proteins, or transgenic pigs expressing a single chain immunoglobulin molecule with specificity for PoEV proteins or other protein that might interfere with protein synthesis or viral assembly may also be produced. Similar transgenes encoding trans-dominant negative regulators of PoEV expression or transgenes encoding competative defective "genomic RNAs" may be used to reduce or eliminate the production of infectious virions. The generation of reagents to suppress the expression of native PoEV loci in pigs, such as, by generation of antisense nucleic acids (e.g. antisense mRNAs), ribozymes or other antiviral reagents may also be developed.

The polynucleotide fragment can be molecularly cloned into a prokaryotic or eukaryotic expression vector using standard techniques and administered to a host. The expression vector is taken up by cells and the polynucleotide fragment of interest expressed, producing protein. Presentation of the protein on cell surface stimulates the host immune system to produce antibodies immunoreactive with said protein as part of a defence mechanism. Thus, expressed protein may be used as a vaccine.

Inactivated vaccines can be produced from PoEV's or cells releasing PoEV. Such infected cells may be generated by natural infection or by transfection of a proviral clone of PoEV. It will be understood that a proviral clone is a molecular clone encoding on at least one antigenic polypeptide of PoEV. After harvesting the virus and/or the infected cells, viruses or infected cells present can be inactivated for example, with formaldehyde, gluteraldehyde, acetylethylenimine or other

suitable agent or process to generate an inactivated vaccine using methods commonly employed in the art. (CVMP Working Party on Immunological Veterinary Medicinal Products (1993). General requirements for the production and control of inactivated mammalian bacterial and viral vaccines for veterinary use). Sub unit vaccines may be prepared from the individual proteins encoded by the gag, pol and env genes. Typically a vaccine would contain env gene products either alone or in combination with gag genes produced by expression in bacteria, yeast or mammlian cell systems.

Proviral clones of PoEV can be engineered to develop single cycle or replication defective viral vectors suitable for vaccination using techniques. Such viral vectors known in the art (e.g. MuLV Murine Leukaemia Retrovirus, Adenovirus and Herpesviruses (Anderson WF. (1992). Human Gene Therapy. *Science* 256, 808-813) may have one or more genes essential for replication deleted, with the missing gene function expressed constitutively or conditionally from a further, different construct which is integrated into the chromosomal DNA of a complementing cell line to the proviral PoEV clone. PoEV virions released from the cell line may infect secondary target cells in the vaccinee but not produce further infectious virus particles. For instance, the polynucleotide sequence encoding the reverse transcriptase domain of *pol* can be deleted from the proviral PoEV clone and the reverse transcriptase domain of *pol* integrated into the complementing cell line.

It will be understood that the polynucleotides; polypeptides; PoEV free cells, tissues and/or organs encompassed

by the present invention could be used in therapy, diagnosis, and/or methods of treatment. The polynucleotides; polypeptides; PoEV free cells, tissues and/or organs encompassed by the present invention can also be used in the preparation of medicaments for use in therapy or diagnosis.

The cloning and expression of a recombinant PoEV polynucleotide fragment also facilitates in producing anti-PoEV antibodies and fragments thereof (particularly monoclonal antibodies) and evaluation of in vitro and in vivo biological activity of recombinant PoEV polymerase and/or envelope polypeptides. The antibodies may be employed in diagnostic tests for native PoEV virus.

It will be understood that for the particular PoEV polypeptides embraced herein, natural variations can exist between individuals or between members of the family Suidae (i.e. the pig family). These variations may be demonstrated by (an) amino acid difference(s) in the overall sequence or by deletions, substitutions, insertions, inversions or additions of (an) amino acid(s) in said sequence. All such derivatives showing active polymerase and/or envelope polypeptide physiological and/or immunological activity are included within the scope of the invention. For example, for the purpose of the present invention conservative replacements may be made between amino acids within the following groups:

(I) Alanine, serine, threonine;

(II) Glutamic acid and aspartic acid;

(III) Arginine and leucine;

(IV) Asparagine and glutamine;

(V) Isoleucine, leucine and valine;

(VI) Phenylalanine, tyrosine and tryptophan

Moreover, recombinant DNA technology may be used to prepare nucleic acid sequences encoding the various derivatives outlined above.

As is well known in the art, the degeneracy of the genetic code permits substitution of bases in a codon resulting in a different codon which is still capable of coding for the same amino acid, e.g. the codon for amino acid glutamic acid is both GAT and GAA. Consequently, it is clear that for the expression of polypeptides with the amino acid sequences shown in Figure 1 or fragments thereof, use can be made of a derivative nucleic acid sequence with such an alternative codon composition different from the nucleic acid sequence shown in said Figure 1.

Furthermore, fragments derived from the PoEV core, polymerase and/or envelope polypeptides as depicted in Figure 3, which still display PoEV virus core polypeptide, polymerase and/or envelope polypeptide properties, or fragments derived from the nucleic acid sequence encoding the virus core polypeptides, polymerase and/or envelope polypeptides or derived from the nucleotide sequence depicted in Figures 1,2,3 and/or 4encoding fragments of said virus core polypeptide, polymerase and/or envelope polypeptides are also included of the present invention. Naturally, the skilled addressee will appreciate within the ambit that the said fragments should substantially retain the physiological and/or immunological properties of the *GAG*, *POL* and/or *ENV* polypeptides.

The PoEV polynucleotide fragment of the present invention

is preferably linked to regulatory control sequences. Such control sequences may comprise promoters, operators, inducers, enhancers, ribosome binding sites, terminators etc. Suitable control sequences for a given host may be selected by those of ordinary skill in the art.

A polynucleotide fragment according to the present invention can be ligated to various expression controlling sequences, resulting in a so called recombinant nucleic acid molecule. Thus, the present invention also includes an expression vector containing an expressible PoEV nucleic acid molecule. The recombinant PoEV nucleic acid molecule can then be used for the transformation of a suitable host. Such hybrid molecules are preferably derived from, for example, plasmids or from nucleic acid sequences present in bacteriophages or viruses and are termed vector molecules.

Specific vectors which can be used to clone nucleic acid sequences according to the invention are known in the art (e.g. Rodriguez, R.L. and Denhadt, D.T., Edit., Vectors: a survey of molecular cloning vectors and their uses, Butterworths, 1988).

The methods to be used for the construction of a recombinant nucleic acid molecule according to the invention are known to those of ordinary skill in the art and are *inter alia* set forth in Sambrook, *et al*. (Molecular Cloning: a laboratory manual Cold Spring Harbour Laboratory, 1989).

The present invention also relates to a transformed cell containing the PoEV polynucleotide fragment in an expressible form. "Transformation", as used herein, refers to the introduction of a heterologous polynucleotide fragment into a

host cell. The method used may be any known in the art, for example, direct uptake, transfection transduction or electro poration (Current Protocols in Molecular Biology, 1995. John Wiley and Sons Inc). The heterologous polynucleotide fragment may be maintained through autonomous replication or alternatively, may be integrated into the host genome. The recombinant nucleic acid molecules preferably are provided with appropriate control sequences compatible with the designated host which can regulate the expression of the inserted polynucleotide fragment, e.g. tetracycline responsive promoter, thymidine kinase promoter, SV-40 promoter and the like.

Suitable hosts for the expression of recombinant nucleic acid molecules may be prokaryotic or eukaryotic in origin. Hosts suitable for the expression of recombinant nucleic acid molecules may be selected from bacteria, yeast, insect cells and mammalian cells.

Since the biological half life and the degree of glycosylation of recombinant PoEV virus core polypeptide, polymerase and/or envelope polypeptides may be important for use *in vivo*, yeast and baculovirus systems, in which a greater degree of processing and glycosylation occur, are preferred. The yeast strain *Pichia Pastoris* exhibits potential for high level expression of recombinant proteins (Clare et al., 1991). The baculovirus system has been used successfully in the production of type 1 interferons (Smith et al., 1983).

Embodiments of aspects of the present invention will now be described by way of example only which are not intended to be limiting thereof.

Examples Section

Example 1

Preparation of viral RNA

500ml of supernatant derived from exponentially growing porcine kidney cells (PK-15, American Type Culture Collection CCL 33) was clarified by centrifugation of approximately 11,000xg for 10 minutes. Virus was pelleted from the clarified supernatant by centrifugation at approximately 100,000xg for 60 minutes. The supernatant was discarded and the viral pellet retained for the preparation of viral RNA genomes. RNA was prepared from the virus pellet using a Dynabeads (registered trade mark) mRNA Direct kit according to the manufacturer's protocols; A PoEV virus pellet was resuspended in 500μl of TNE (10mM Tris HCl pH8.0, 0.1M NaCl,1mM EDTA) and the virions disrupted by the addition of 2ml of lysis/binding buffer. Dynabeads Oligo$(dT)_{25}$ were conditioned according to the manufacturer's instructions and added to the virus disrupted solution. Viral RNA was allowed to bind to the Dynabead for 10 minutes before the supernatant was removed and the bound RNA was washed three times with washing buffer with LiDS (0.5ml) and twice with washing buffer alone. The RNA was finally resuspended in 25 μl of elution solution. All procedures were performed at ambient temperature. RNase contamination was avoided by the wearing of gloves, observation of sterile technique and treatment of solutions and non-disposable glass and plasticware with diethyl pyrocarbonate (DEPC). The RNA was resuspended in DEPC- treated sterile water.

Example 2

Synthesis of cDNA

cDNA was synthesised from the purified genomic RNA using Great Lengths ™ cDNA amplification reverse transcriptase reagents (Clontech Laboratories Inc.) following the manufacturer's instructions. The RNA was primed with both oligo(dT) and random hexamers to maximise synthesis.

The Great Lengths cDNA synthesis protocol is based on a modified Gubler and Hoffman (1983) protocol for generating complementary DNA libraries and essentially consists of first-strand synthesis, second strand synthesis, adaptor ligation, and size fractionaction.

First strand synthesis: lock-docking primers anneal to the beginning of the poly-A tail of the RNA due to the presence of A, C or a residue at the 3'-end of the primer. This increases the efficiency of cDNA synthesis of eliminating unnecessary reverse transcription of long stretches of poly-A. In addition, the reverse transcriptase used is MMLV (RNase H$^-$) which gives consistently better yields than do wild-type MMLV or AMV reverse transcriptase.

Second strand synthesis: the ratio of DNA polymerase I for RNase H has been optimised to increase the efficiency of the second strand synthesis and to minimize priming by hair pin loop formation. Following secoond-strand synthesis, the ds cDNA is treated with T4 DNA polymerase to create blunt ends.

Adaptor ligation: the cDNA is ligated to a specially designed adaptor that has a pre-existing EcoRI "sticky end". The use of this adaptor, instead of a linker, eliminates the need to methylate and the EcoRI - digest the cDNA, and thus leaves internal EcoRI, sites intact. The adaptor is 5'-phosphorylated at the blunt end to allow efficient ligation to the blunt-ended cDNA.

Size fractionation: the ds cDNA is phosphorylated at the EcoRI sites and size-fractionated to remove unligated adaptors and unincorporated nucleotides. The resulting cDNA is ready for cloning into a suitable EcoRI-digested vector.

Example 3

Molecular cloning of cDNA

The size fractionated fragment was ligated with EcoR I- digested pZErO™ -1 plasmid vector DNA (Invitrogen Corporation, San Diego, U.S.). The ligation mix was used to transform competent TOP10F'cells and these were plated onto L-Agar containing zeocin following the manufacturer's instructions (Zero Background™ cloning kit - Invitrogen). Several of the resulting zeocin resistant colonies were amplified in L-Broth containing zeocin and the plasmid DNA was purified by alkaline lysis (Maniatis *et al.*, 1982).

The plasmid DNA was digested to completion with the endonuclease EcoR I and the resulting DNA fragments were separated by electrophoresis through an 1.0% agarose gel (Maniatis *et al.*, 1982), in order to check that a fragment in the

predicted size fractionated size range had been cloned. A clone identified as pPoEV was used in further experimentation.

Example 4

DNA sequence analysis.

pPoEV plasmid DNA was purified according to common techniques (Sambrook et al, 1989) and sequenced using an ABI automated sequencer. Overlapping sequencing primers from both strands of the molecular clone were used to determine the nucleotide sequence.

The first sequence obtained is shown in Figure 1. This sequence was identified as encoding two ORFs of 924 (nucleotides 23-2793) and 218 (nucleotides 2642-3297) amino acids, relating to the *pol* and *env* genes respectively. This sequence was revised and updated to the second sequence as shown in Figure 2. This second sequence was identified as encoding three ORFs of 516 (nucleotides 576-2126), 1186 (nucleotides 2143-5733) and 656 (nucleotides 5606-7576) amino acids, encoding the PoEV *gag*, *pol* and *env* genes respectively. This second sequence has since been revised and updated to the sequnce shown in Figure 3. This third sequence was identified as encoding three ORFs of 524 (nucleotides 588-2162), 1194 (nucleotides 2163-5747) and 656 (nucleotides 5620-7590) amino acids, encoding the PoEV *gag*, *pol* and *env* genes respectively.

The differences in the disclosed segeunces is reflected by improvements in carrying out and analysing the sequence obtained. However, there is 100% identity at the nucleic acid level, between positions 21-2681 of the first sequence and positions 2972-5653 of the third sequence. Overall there is a 70.5%

identity in the entire 3310 bp of the first sequence with a corresponding portion of the third sequence.

There are only 3 base changes between the second sequence and the third sequence. These are as follows:

base no. (from Figure 2)	change
2121	insertion of a "G"
2157	insertion of a "G"
5902	"R" to an "A"
7700	"M" to an "A"

The changes at base nos. 5902 and 7700 do not effect the corresponding amino acid sequence. However, the changes at positions 2121 and 2157 alter the amino acid sequence at the end of GAG and the begining of POL. For GAG the final amino acid "S" have now been replaced by "VLALEEDKD". The total product size is now 524 amino acids. For POL, the first five amino acids "RLGET" have been deleted and replaced by "GRR". The total product size is now 1194 amino acids.

Similarities were observed between pPoEV and the majority of retroviruses determined by using alogrithims from DNASTAR Inc. Lasergene software (DNASTAR). The similarities were closest with gibbon ape leukaemia virus (GaLV) in the polymerase (*pol*) regions of the pro-virus at 68.5%, in the virus core *(gag)* region, 59.2% and in the envelope *(env)* region, 39.3% The nucleotide sequence and major ORFs of the pPoEV insert are shown in Figure 3. The largest ORF (nucleotides 2163-5747) encodes the polymerase polypeptide and the smaller ORFs (nucleotides 588-2162 and 5620-7590) encode the core and envelope polypeptides respectively.

Example 5

Purification of cellular DNA from cultured cells, tissues and blood.

Cultured cells

Cells were maintained in culture and approximately 5 x 10^7 cells were harvested for DNA preparation. The cells were pelleted by centrifugation resuspended in phosphate-buffered saline, re-centrifuged at 1000g for 2 minutes and the supernatant was discarded.

Porcine tissues

Porcine tissue samples were frozen in liquid nitrogen and powdered by grinding in a mortar or between metal foil. The samples were resuspended in 5ml of extraction buffer consisting of 0.025M EDTA (pH 8.0), 0.01MTris.Cl pH 8.0, 0.5% SDS 20μg/ml RNAse and 100μg/ml proteinase K (Maniatis *et al.*, 1982).

Porcine blood

A buffy coat was prepared from the blood samples. 20ml samples were centrifuged at 1000g for 15 minutes. The buffy coat was resuspended in buffer and the samples centrifuged at 1000g for 15 minutes. The process was repeated one further time. The sample was mixed with 5ml (3x volume) of extraction buffer (Maniatis *et al.*, 1982).

Purification

The samples (i.e. cultured cells, porcine tissue or porcine blood cells) in proteinase K-extraction buffer containing 20μg/ml RNAse

and 100μg/ml proteinase K were digested for approximately 24 hours at 37°C. The deproteinised DNA was extracted twice with phenol and twice with phenol chloroform and finally precipitated by ethanol in the presence of ammonium acetate. The DNA was recovered by centrifugation at 3000g for 30 minutes and the supernatant discarded (Maniatis *et al.*, 1982). The pellet was washed in 70% ethanol and allowed to air dry for approximately 1 hour. The DNA was allowed to re-dissolve in Tris EDTA (TE) buffer and the purity and concentration of the DNA was assessed by spectrophotometry (Maniatis *et al.*, 1982).

Example 6

Southern blot analysis of porcine tissue and cells

In order to demonstrate that the molecularly cloned DNA comprising the insert from PoEV was derived from the PK-15 cell line (American Type Culture Collection CCL33), the DNA was hybridised against cellular DNAs and its ability to detect proviral DNA was examined.

DNA purified from pPoEV was radioactively labelled and used to probe a Southern blot of endonuclease digested DNAs derived from PK-15 cells .

The DNAs probed were as follows :

a) Copy number controls of pPoEV DNA linearized by digestion with EcoRI. One copy per haploid cell genome was estimated to be 6.84pg. The control was present at an estimated copy number of 1, 5 and 10 copies.

b) PK-15 DNA.

c) Negative control HeLa (American Type Culture Collection

CCL2) DNA derived from a human adenocarcinoma cell line harbouring human papillomavirus type 18 DNA.

d) Negative control SP20 (European Collection of Animal Cell Cultures 85072401) DNA derived from a murine myeloma cell line harbouring a xenotropic MuLV retrovirus.

A hybridisation signal was observed in only the PK-15 porcine DNA. No signal was detected in either the negative human or murine DNAs. The PK-15 DNA contained more than 10 copies per cell with an estimated copy number of 20. The sizes of the three major EcoRI- endonuclease digested DNA fragments were approximately 3.8kb, 1.8kb and 0.6kb. The sizes of relevant fragments detected in the recombinant pPoEV were comparable.

There are, as expected, a number of fragments common to the genomic DNA of PK-15 and pPoEV DNA and there is agreement between the patterns observed in both DNAs digested with XhoI, BamHI and HindIII. However, there are additional fragments obtained on digestion of pPoEV DNA by a number of endonucleases.

pPoEV sequences were also detected in swine testes (American Type Culture Collection CRL 1746) and primary porcine kidney cells (Central Veterinary Laboratory batch C04495) but not in hamster CHOK1 (American Type Culture Collection CCL61) or murine NS0 myeloma cells (European Collection of Animal Cell Cultures 85110503).

In order to demonstrate that the molecularly cloned DNA comprising the insert from pPoEV could detect sequences in porcine cells and tissues in addition to PK-15 the pPoEV DNA was hybridised against cellular DNA from tissues derived from pigs and its ability to detect proviral DNA was examined (Maniatis *et al.*, 1982).

The DNA purified from pPoEV was radioactively labelled and used to probe a Southern blot of endonuclease digested DNAs derived from pig organs including liver, kidney, heart and blood.

The DNAs probed were as follows :

a) Copy number controls of pPoEV DNA linearized by digestion with EcoRI. One copy per haploid cell genome was estimated to be 6.84pg. The control was present at an estimated copy number of 5,10, 20 and 50 copies.

b) DNA purified from the porcine tissues digested with EcoRI.

A hybridisation signal was observed in all the porcine DNAs.

The DNAs contained less than 5 copies per cell. There were approximately eight distinct bands in each DNA. The sizes of the three major endonuclease digested DNA fragments were approximately 3.8kb, 1.8kb and 0.6kb.

Example 7

Polymerase Chain Reaction (PCR) Amplifications

Oligonucleotides were selected from the PoEV genome.

The upstream primer was 5'-GGA AGT GGA CTT CAC TGA G-3'.

The downstream primer was 5'-CTT TCC ACC CCG AAT CGG -3'.

The PCR was performed as described by Saiki *et al* (1987). One 1μl of 100ng/μl template DNA was added to a 49μl reaction mixture containing 200μM of dATP, dCTP, dGTP, dTTP, 30pmol of both primers from the pair described above, 1unit of DNA polymerase and 5μl of reaction buffer. The reaction buffer contained 200mM Tris-HCl pH 8.4, 500mM potassium chloride and 15mM magnesium chloride, ultrapure water. The solution was overlaid with two drops of mineral oil to prevent evaporation. Thirty five cycles of amplification were performed using a Perkin Elmer Cetus thermal cycler. Each cycle consisted of 1 minute. at 95^{0}C to denature the DNA, 1 minute. at 53°C to anneal the primers to the template and 1 minute. at 72^{0}C for primer extension. After the last cycle a further incubation for 10 minutes. at 72^{0}C was performed to allow extension of any partially completed product. On completion of the amplification, 10μl of the reaction mixture was electrophoresed through a 5 per cent acrylamide gel. The DNA was visualised by staining with ethidium bromide and exposure to ultraviolet light (320nm).

The PCR reaction amplified a sequence of approximately 787bp from pPoEV and from porcine cells as expected indicating that the assay detected the PoEV proviral DNA. There was no specific amplification of the expected sequence in cells of non-porcine origin and therefore, the PCR reaction and recombinant clone can be used as a specific and sensitive diagnostic tool for detection of PoEV.

Two further digonucleotides were designed against the 3'end of the *pol* gene and s' end of the *gag* gene respectively.

The 3' *pol* oligionucleotide was 5'-GAT GGC TCT CCT GCC CTT TG-3'

The 5' *gag* oligionucleotide was 5'-CGA TGG AGG CGA AGC TTA AGG-3'

The above oligionucleotide were also used in in PCR reactions according to the conditions described above, with the exceptions that the annealing temperature was 58° and 30 cycles of replication were carried out. The PCR reaction amplified a sequence of approximately 468bp from pPoEV and from porine cells.

Example 8

Production of PoEV polypeptide in *Escherichia coli*.

The open reading frame (ORF) encoding the *pol* peptide was isolated from the pPoEV clone and molecularly cloned into the plasmid pGEX-4T-1 (Pharmacia Ltd.) for expression.

Two ml cultures of *E. coli* transformed with various expression constructs were grown with shaking at 37°C to late log phase

($O.D._{600nm}$ of 0.6) and induced by the addition of IPTG to 0.1 mM. Induced cultures were then incubated for a further 2 hours after which the bacteria were collected by centrifugation. The bacterial pellet was lysed by boiling in SDS-PAGE sample buffer and the protein profile of the induced bacteria was analysed on a 12% acrylamide gel (Laemmli, 1970) followed by staining with coomassie brilliant blue dye.

Example 9

Isolation and partial sequencing of Raji clone

The aim of the study was to determine whether the human cell line "Raji" was susceptible to infection with the PoEV present in porcine kidney cells (PK15). In order to test the capacity of the virus for xenotropism, PK15 cells were co-cultured with the B lymphoblastoid (Raji) cell line over 20 passages.

The culture system utilised direct culture and transwells, which separated the human and porcine cells, but permitted viruses to pass through the separating membrane. After every fifth passage, supernatants from the human cell lines are tested for the presence of retrovirus by reverse transcriptase assay.

Cell cultures

Porcine kidney (PK15) cells (ATCC CCL 33) were used as the source of PoEV. The human cells used for co-cultivation with PK15 cells were the lymphoblast-like Burkitts lymphoma Raji (ATCC CCL 86) cell line. This cell line does not harbour endogenous

retroviruses and lacks reverse transcriptase activity when tested by the present inventors.

Co-cultivation

Raji cells were co-cultivated directly with PK15 cells in duplicate 80cm^2 flasks and exposed to the PK15 cells throughout the 20 passage culture period. The cells were passaged twice per week and PK15 cells added as necessary from a stock culture. At every fifth passage a sample of Raji cells was removed from the co-culture, washed and cultured for 3-4 days. Supernatant was then harvested and tested for presence of retrovirus by reverse transcriptase (RT) assay.

RESULTS

The presence of reverse transcriptase activity with a preference for the Mn^{2+} cation in the supernatant from detector cell cultures is suggestive of infection by porcine retrovirus. Reverse transcriptase activity with preference for the Mn^{2+} template was not detected in the duplicate co-cultivated test cultures at passage 5 but was detected at passages 10, 15 and 20. No significant RT activity was detected in the negative control cultures. RT activity with preference for the Mn^{2+} template was detected in positive control cultures at passage 5 and 20.
An infected raji culture was diluted to single cells, and then a selection of cells cultured separately such that each culture originated from one cell. Each culture was tested by reverse-transcriptase assay.

Genomic DNA was made from an RT-positive clone as described in example 5 -purification. The PoEV ENV region was amplified by PCR as described below and the product molecularly cloned into pMOS blue T-vector (Amersham). This molecular clone was then sequenced (Fig. 4).

PCR

Oligonucleotides were selected from the PoEV genome.

The upstream primer was 5'-GAT GGC TCT CCT GCC CTT TG -3'
5' base position: 5240
The downstream primer was 5'-CCA CAG TCG TAC ACC ACG -3'
5' base position: 8144

Expected product size: 2904bp

Approx. 1 μg of genomic raji clone DNA was added to a 50 μl reaction mixture containing 200 μM of dATP, dCTP, dGTP, dTTP, 30pM each primer detailed above, 1u Taq DNA polymerase and 5μl reaction buffer. The reaction buffer contained 200mM Tris.Cl pH 8.4, 500mM potassium chloride, 15mM magnesium chloride and ultrapure water. The solution was overlaid with two drops of mineral oil to prevent evaporation. Thirty cycles of amplification was performed followed by an elongated extension reaction of 60min. at 72°C.

The cycles consisted of:

95°C 1 min.

56°C 1 min.

72°C 2 min.

The PCR product was visualised as described in example 7. Product size: ~3Kb.

CLONING

The PCR product was molecularly cloned into pMOS-Blue T-vector as directed by the manufacturer (pMOS-Blue T-vector kit - Amersham).

20 transformed colonies (clones) were picked and added to 5mls L-broth containing 50 μg/ml ampicillin. The cultures were grown shaking at 37°C overnight. Plasmid DNA was isolated from each clone using the perfect prep plasmid isolation kit as directed by the manufacturer (5 Prime - 3 Prime Inc. Boulder, CO, USA).

Plasmid DNA was digested to completion with the endonucleases EcoRI and HindIII and the products visualised on an ethidium bromide-stained 1% agarose gel. A clone (raji env clone)showing the same banding pattern as that predicted for 'PK15 cell line derived PoEV', was selected for sequencing.

SEQUENCING

Raji env clone plasmid DNA prepared above was sequenced using an ABI automated sequencer, and the commercially availableT7 sequencing primer.

The entire *env* gene region of the "Raji" was sequnced (see Figure 4) and discovered to have substantial sequence identity at both the nucleic acid and amino acid levels (98.9% and 96.3% respectively) with the PoEV sequence from PK-15.

Example 10

Phylogenetic analysis

Phylogenetic analysis was performed using the PHYLIP package. Sequence distances were calculated using the PROTDIST program (Dayhoff matrix) and a neighbour-joining unrooted phylogenetic tree reconstructed using the NEIGHBOUR program.

Bootstrapping was performed using 200 replicates of the *pol* alignment, created using the SEQBOOT program and a consensus tree was obtained using the CONSENSE program (see Figure J). The bootstrap percentages are indicated at the branch fork, with missing values equal to 100%. The data indicate that PoEV is closely related to but distinct from the type-C oncovirus typified by gibbon, murine and feline leukaemia viruses.

A phylogenetic tree was constructed from the *pol* alignment using the maximum likliehood algorithm (Dayhoff matrix). This tree differed from the *pol* NJ tree only in the placement of the BaEV lineage in relation to other mammalian type C viruses and corresponded to a low bootstrap for the BaEV fork observed in the NJ tree.

Example 11

Analysis of the LTR and adjacent region

The long terminal repeat (LTR) is a reiterated sequence at each end of the provirus that contains the enhancer and promoter governing transcription of the provirus as well as sequences required for reverse transcription of the RNA genome and integration of the proviral DNA. Three recognised domains of the LTR are identifiable, U3, R and U5 with the LTR being delineated by inverse repeats AATGAAAGG and CCTTTCATT at the 5′ and 3′ ends of U3 and U5 respectively.

LTR Domain	PoEV Genome Sequence in accordance with Figure 3	Length bp
U3	7638-8106	469
R*	8107-8188,1-61	82
U5	62-143	82

*The position of the R is defined here by similarity to the 3′end of the MuLV LTR and is compatible with the observed location of a cap site approximatelty 24 bp downstream of the TATA box.

The U3 region contans multiple potential transcription sites as shown in Figure 6. Most of the U3 region shows little or no homology to other mammalian type-C retroviruses which show conserved sites or repeat elements. However, there is homology to other mammaliann type-C viruses towards the 3′end of the U3 & region and into R and U5. Amongst the potential transcription factor sites are those for the following:

LyF-1 is a transcriptional regulator that interacts with a novel class of promoters for lymphocyte-specific-genes (Lo et al 1991).

E47 is the prototype member of a new family of tissue specific enhancer proteins that have been shown to bind to the enhancer of murine leukaemia virus.

ETS-1 is a transcription factor primarily expressed in the haematopoietic lineage.

The LTR contains direct repeats at 80006-8062 and 8045-8101 which together contain three potential CCAATT boxes. A potential TATA box is located at position 8129-8144.

The R region contains a PADS (Poly A downstream element) and consensus polyadenylation signal (AATAAA).

The primer binding site (PBS) of PoEV is glycine(2) tRNA which has not reported for any exogenous retrovirus.

References

Clare JJ, Rayment FB, Ballantine SP, Sreekrishna K and Romanos MA. (1991). High Level expression of tetanus toxin fragment C in *Pichis pastoris* strains containing multiple tandem integrations of the gene. *Bioltechnology, 9,* 455-460

Derynck R, Singh A and Goeddel DV. (1983). Expression of the human interfron-γ in yeast. *Nucleic Acids Res.*, 11, 1819-1837.

DNASTAR. (1994). Lasergene Biocomputing Software for Windows. User's Guide.

Invitrogen. Version A. Zero Background™ Cloning Kit Catalog no K2500-01.

Laemmli UK. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4N. *Nature*, 227, 680-685.

Lieber MM, Sherr CJ. Benveniste RE and Torado GJ. (1975). Biologic and immunologic properties of porcine type C viruses. *Virology* 66, 616-619.

Lo K, LAndau NR, Smale ST. *Mol. Cell. Biol.* 11:5229-5243(1991)

Maniatis T, Fritsch EF and Sambrook J. (1982) *Molecular Cloning: A Laoratory Manual,* Cold pring Harbour Laboratory, Cold Spring Harbour, NY.

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT Mullis KB and Erlich HA. (1987). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239, 487-491.

Sambrook J Fritsch EF, and Maniatis T. (1989). *Molecular Cloning a Laboratory Manual, 2nd ed.* Cold Spring Harbour Laboratory, Cold Spring Harbour New York.

Smith GE, Summers MD and Fraser MJ. (1983). Production of human beta interferon in insect cells infected with a baculovirus vector. *Mol. Cell. Biol.*, 3, 2156-2165.

Stranstrom H, Verjalainen P, Meoning V Hunsmann G, Schwarz H. and Schafer W. (1974). C- type particles produced by a permanent cell line from a leukemic pig. 1 Origin and properties of host cells and some evidence for the occurence of C-Type like particles *Virology* 57, 175-178.

Todaro GJ, Benveniste RE, Lieber MM and Sherr CJ. (1974). Characterizaation of a type C virus released from the porcine cell line PK (15). *Virology* 58, 65-74.

Claims

1. An isolated polynucleotide fragment:

   (a) encoding at least one porcine retrovirus (PoEV) expression product;
   (b) encoding a physiologically active and/or immunogenic derivative of said expression product; or
   (c) which is complementary to a polynucleotide sequence as defined in (a) or (b).

2. An isolated polynucleotide fragment according to claim 1:

   (a) encoding the polymerase (POL) polypeptide;
   (b) encoding a physiologically active and/or immunogenic derivative of a polypeptide as described in (a); or
   (c) which is complementary to a polynucleotide sequence as defined in (a) or (b).

3. An isolated polynucleotide fragment according to claim 1:

   (a) encoding the virion core polypeptide (GAG) and/or envelope polypeptide (ENV);
   (b) encoding a physiologically active and/or immunogenic derivative of a polypeptide as described in (a); or
   (c) which is complementary to a polynucleotide sequence as defined in (a) or (b).

4. An isolated polynucleotide fragment according to claim 1:

(a) encoding the virion core polypeptide (GAG), polymerase (POL) and envelope polypeptide (ENV);
(b) encoding a physiologically active and/or immunogenic derivative of a polypeptide as described in (a); or
(c) which is complementary to a polynucleotide sequence as defined in (a) or (b).

5. An isolated polynucleotide fragment according to any one of claims 1 to 4 wherein the polynucleotide fragments is a deoxyribose nucleic acid (DNA) fragment.

6. An isolated polynucleotide fragment according to any preceding claim encoding:

(a) said at least one polypeptide having an amino acid sequence which is shown in Figures 3 or 4;
(b) encoding a polypeptide which is a physiologically active and/or immunogenic derivative of at least one of the polypeptides defined in (a); or
(c) which is complementary to a polynucleotide sequence as defined in (a) or (b).

7. An isolated polynucleotide fragment according to any preceding claim;

(a) comprising at least one of the ORFs shown in Figures 1,2,3 or 4 or comprising a corresponding RNA sequence;

(b) comprising a sequence having substantial nucleotide sequence identity with a sequence as described in (a) above; or

(c) comprising a sequence which is complementary to a sequence as described in (a) or (b) above.

8. A recombinant nucleic acid molecule comprising a polynucleotide fragment according to any one of claims 1 to 7.

9. A recombinant nucleic acid molecule according to claim 8 wherein the recombinant nucleic acid molecule comprises regulatory control sequences operably linked to said polynucleotide fragment for controlling expression of said polynucleotide fragment.

10. A vector comprising a recombinant nucleic acid molecule according to either of claims 8 or 9.

11. A vector according to claim 10 which is a virus or a plasmid.

12. A prokaryotic or eukaryotic host cell transformed by a polynucleotide fragment, recombinant nucleic acid molecule, or vector according to any of claims 1 to 11.

13. A recombinant PoEV polypeptide or derivative thereof displaying POL PoEV physiological and/or immunogenic activity.

14. A recombinant PoEV polypeptide or derivative thereof displaying GAG and/or ENV PoEV physiological and/or immunogenic activity.

15. A recombinant PoEV polypeptide or derivative thereof displaying GAG, POL and ENV PoEV physiological and/or immunogenic activity.

16. A recombinant PoEV polypeptide according to any one of claims 13 to 15 comprising a sequence as shown in Figures 3 or 4, or functionally active derivative thereof.

17. A vaccine comprising a recombinant PoEV polypeptide according to any one of claims 13 to 16, or an inactivated PoEV virus and a pharmaceutically acceptable carrier.

18. An antibody or fragment thereof capable of binding to a polypeptide or fragment according to any one of claims 13 or 16.

19. A polynucleotide primer which is PoEV specific.

20. A polynucleotide probe which is capable of specifically hybridising to a PoEV polynucleotide sequence.

21. A probe or a primer according to claims 19 or 20 which has substantial nucleotide sequence identity with a strand of the molecule depicted in Figures 1,2,3 or 4 or a strand complementary therewith, with a corresponding RNA molecule, or with a part of such a molecule.

22. A PoEV detection kit comprising a polynucleotide primer or probe according to any of claims 19 to 21.

23. Use of a PoEV specific polynucleotide in the detection of PoEV in a sample.

24. Use of a PoEV specific polynucleotide in a PCR for the detection of PoEV in a sample.

25. A pig modified so as to not express an infectious PoEV capable of infecting humans.

26. Cells, tissues or organs obtainable from a pig accoding to claim 25.

27. Use of a recombinant PoEV polypeptide according to any one of claims 13 to 16 in the preparation of a vaccine.

28. Use of a polynucleotide primer or probe according to any of claims 19 to 21 in the preparation of a detection kit capable of detecting the presence of PoEV nucleic acid in a sample.

29. Use of a polynucleotide; polypeptide; cells, tissues or organs according to any one of claims 1 to 7, 13 to 16 or 26 in therapy or diagnosis.

30. A polynucleotide; polypeptide; cells, tissues or organs according to any one of claims 1 to 7, 13 to 16 or 26 in the preparation of a medicament for use in therapy or diagnosis.

31. The invention substantially as hereinbefore described.

Figure 1

```
   1 GAATTCGCGGCCGCGTCGACAGATGCCTTCTTCTGCCTGAGATTACACCCCACTAGCCAA   60

  61 CCACTTTTTGCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAGCTCACCTGG  120

 121 ACCCGACTGCCCCAAGGGTTCAAGAACTCCCCGACCATCTTTGACGAAGCCCTACACAGG  180

 181 GACCTGGCCAACTTCAGGATCCAACACCCTCAGGTGACCCTCCTCCAGTACGTGGATGAC  240

 241 CTGCTTCTGGCGGGAGCCACCAAACAGGACTGCTTAGAAGGTACGAAGGCACTACTGCTG  300

 301 GAATTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAGGCCCAGATTTGCAGGAGAGAG  360

 361 GTAACATACTTGGGGTACAGTTTGCGGGGCGGGCAGCGATGGCTGACGGAGGCACGGAAG  420

 421 AAAACTGTAGTCCAGATACCGGCCCCAACCACAGCCAAACAAGTGAGAGAGTTTTTGGGG  480

 481 ACAGCTGGATTTTGCAGACTGTGGATCCCGGGGTTTGCGACCTTAGCAGCCCCACTCTAC  540

 541 CCGCTAACCAAAGAAAAAGGGGGATTCTCCTGGGCTCCTGAGCACCAGAAGGCATTTGAT  600

 601 GCTATCAAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACTAAACCC  660

 661 TTTACCCTTTATGTGGATGAGCGTAAGGGAGTAGCCCGAGGAGTTTTAACCCAAACCCTA  720

 721 GGACCATGGAGGAGACCTGTTGCCTACCTGTCAAAGAAGCTTGATCCTGTAGCCAGTGGT  780

 781 TGGCCCGTATGTCTGAAGGCTATCGCAGCTGTGGCCATACTGGTCAAGGACGCTGACAAA  840

 841 TTGACTTTGGGACAGAATATAACTGTAATAGCCCCCCATGCATTGGAGAACATCGTTCGG  900

 901 CAGCCCCCAGACCGATGGATGACCAACGCCCGCATGACCCACTATCAAAGCCTGCTTCTC  960

 961 ACAGAGAGGGTCACTTTCGCTCCACCAGCCGCTCTCAACCCTGCCACTCTTCTGCCTGAA 1020

1021 GAGACTGATGAACCAGTGACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGGTC 1080

1081 CGCAAGGACCTTACAGACATACCGCTGACTGGAGAAGTGCTAACCTGGTTCACTGACGGA 1140

1141 AGCAGCTATGTGGTGGAAGGTAAGAGGATGGCTGGGGCGGCAGTGGTGGACGGGACCCGC 1200

1201 ACGATCTGGGCCAGCAGCCTGCCGGAAGGAACTTCAGCGCAAAAGGCTGAGCTCATGGCC 1260
```

Figure 1 cont.

1261 CTCACGCAAGCTTTGCGGCTGGCCGAAGGGAAATCCATAAACATTTATACGGACAGCAGG 1320

1321 TATGCCTTTGCGACTGCACACGTACACGGGGCCATCTATAAACAAAGGGGGTTGCTTACC 1380

1381 TCAGCAGGGAGGGAAATAAAGAACAAAGAGGAAATTCTAAGCCTATTAGAAGCCTTACAT 1440

1441 TTGCCAAAAAGGCTAGCTATTATACACTGTCCTGGACATCAGAAAGCCAAAGATCTCATA 1500

1501 TCTAGAGGGAACCAGATGGCTGACCGGGTTGCCAAGCAGGCAGCCCAGGCTGTTAACCTT 1560

1561 CTGCCTATAATAGAAACGCCCAAAGCCCCAGAACCCAGACGACAGTACACCCTAGAAGAC 1620

1621 TGGCAAGAGATAAAAAAGATAGACCAGTTCTCTGAGACTCCGGAGGGGACCTGCTATACC 1680

1681 TCATATGGGAAGGAAATCCTGCCCCACAAAGAAGGGTTAGAATATGTCCAACAGATACAT 1740

1741 CGTCTAACCCACCTAGGAACTAAACACCTGCAGCAGTTGGTCAGAACATCCCCTTATCAT 1800

1801 GTTCTGAGGCTACCAGGAGTGGCTGACTCGGTGGTCAAACATTGTGTGCCCTGCCAGCTG 1860

1861 GTTAATGCTAATCCTTCCAGAATACCTCCAGGAAAGAGACTAAGGGGAAGCCACCCAGGC 1920

1921 GCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAATACGGAAACAAATATCTA 1980

1981 TTGGTTTTTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCCTACTAAGAAAGAGACT 2040

2041 TCAACCGTGGTGGCTAAGAAAATACTGGAGGAAATTTTTCCAAGATTTGGAATACCTAAG 2100

2101 GTAATAGGGTCAGACAATGGTCCAGCTTTCGTTGCCCAGGTAAGTCAGGGACTGGCCAAG 2160

2161 ATATTGGGGATTGATTGGAAACTGCATTGTGCATACAGACCCCAAAGCTCAGGACAGGTA 2220

2221 GAGAGGATGAATAGAACCATTAAAGAGACCCTTACCAAATTGACCACAGAGACTGGCATT 2280

2281 AATGATTGGATGGCTCTCCTGCCCTTTGTGCTTTTTAGGGTGAGGAACACCCCTGGACAG 2340

2341 TTTGGGCTGACCCCCTATGAATTGCTCTACGGGGGACCCCCCCCGTTGGCAGAAATTGCC 2400

2401 TTTGCACATAGTGCTGATGTGCTGCTTTCCCAGCCTTTGTTCTCTAGGCTCAAGGCGCTC 2460

2461 GAGTGGGTGAGGCAGCGAGCGTGGAAGCAGCTCCGGGAGGCCTACTCAGGAGGAGACTTG 2520

Figure 1 cont.

```
2521 CAAGTTCCACATCGCTTCCAAGTTGGAGATTCAGTCTATGTTAGACGCCACCGTGCAGGA 2580
2581 AACCTCGAGACTCGGTGGAAGGGACCTTATCTCGTACTTTTGACCACACCAACGGCTGTG 2640
2641 AAAGTCGAAGGAATCCCCACCTGGATCCATGCATCCCACGTTAAGCCGGCGCCACCTCCC 2700
2701 GATTCGGGGTGGAAAGCCGAAAAGACTGAAAATCCCCTTAAGCTTCGCCTCCATCGCGTG 2760
2761 GTTCCTTACTCTGTCAATAACTCCTCAAGTTAATGGTAAACGCCTTGTGGACAGCCCGAA 2820
2821 CTCCCATAAACCCTTATCTCTCACCTGGTTACTTACTGACTCCGGTACAGGTATTAATAT 2880
2881 TAACAGCACTCAAGGGGAGGCTCCCTTGGGGACCTGGTGGCCTGAATTATATGTCTGCCT 2940
2941 TCGATCAGTAATCCCTGGTCTCAATGACCAGGCCACACCCCCCGATGTACTCCGTGCTTA 3000
3001 CGGGTTTTACGTTTGCCCAGGACCCCCAAATAATGAAGAATATTGTGGAAATCCTCAGGA 3060
3061 TTTCCTTTGCAAGCAATGGAGCTGCATAACTTCTAATGATGGGAATTGGAAATGGCCAGT 3120
3121 CTCTCAGCAAGACAGAGTAAGTTACTCTTTTGTTAACAATCCTACCAGTTATAATCAATT 3180
3181 TAATTATGGCCATGGGAGATGGAAAGATTGGCAACAGCGGGTACAAAAAGATGTACGAAA 3240
3241 TAAGCAAATAAGCTGTCATTCGTTAGACCTAGATTACTTAAAAATAAGTTTCACTAAAAA 3300
3301 AAAAAAAAAAAAAAAAAAAA 3320
```

Figure 2

```
   1 TGTGGGCCCCAGCGCGCTTGGAATAAAAATCCTCTTGCTGTTTGCATCAAGACCGCTTCT   60
  61 CGTGAGTGATTTGGGGTGTCGCCTCTTCCGAGCCCGGACGAGGGGGATTGTTCTTTTACT  120
 121 GGCCTTTCATTTGGTGCGTTGGCCGGGAAATCCTGCGACCACCCCTTACACCCGAGAACC  180
 181 GACTTGGAGGTAAAGGGATCCCCTTTGGAACGTGTGTGTGTGTCGGCCGGCGTCTCTGTT  240
 241 CTGAGTGTCTGTTTTCGGTGATGCGCGCTTTCGGTTTGCAGCTGTCCTCTCAGACCGTAA  300
 301 GGACTGGAGGACTGTGATCAGCAGACGTGCTAGGAGGATCACAGGCTGCCACCCTGGGGG  360
 361 ACGCCCCGGGAGGTGGGGAGAGCCAGGGACGCCTGGTGGTCTCCTACTGTCGGTCAGAGG  420
 421 ACCGAGTTCTGTTGTTGAAGCGAAAGCTTCCCCCTCCGCGGCCGTCCGACTCTTTTGCCT  480
 481 GCTTGTGGAAGACGCGGACGGGTCGCGTGTGTCTGGATCTGTTGGTTTCTGTCTCGTGTG  540
 541 TCTTTGTCTTGTGCGTCCTTGTCTACAGTTTTAATATGGGACAGACAGTGACTACCCCCC  600
 601 TTAGTTTGACTCTCGACCATTGGACTGAAGTTAGATCCAGGGCTCATAATTTGTCAGTTC  660
 661 AGGTTAAGAAGGGACCTTGGCAGACTTTCTGTGCCTCTGAATGGCCAACATTCGATGTTG  720
 721 GATGGCCATCAGAGGGGACCTTTAATTCTGAAATTATCCTGGCTGTTAAGGCAATCATTT  780
 781 TTCAGACTGGACCCGGCTCTCATCCTGATCAGGAGCCCTATATCCTTACGTGGCAAGATT  840
 841 TGGCAGAAGATCCTCCGCCATGGGTTAAACCATGGCTAAATAAACCAAGAAAGCCAGGTC  900
 901 CCCGAATCCTGGCTCTTGGAGAGAAAAACAAACACTCGGCCGAAAAAGTCGAGCCCTCTT  960
 961 CCTCGTATCTACCCCGAGATCGAGGAGCCGCCGACTTGGCCGGAACCCCAACCTGTTCCC 1020
1021 CCACCCCCTTATCCAGCACAGGGTGCTGTGAGGGGACCTCTGCCCCTCCTGGAGCTCCGG 1080
1081 TGGTGGAGGGACCTGCTGCCGGGACTCGGAGCCGGAGAGGCGCCACCCCGGAGCGGACAG 1140
1141 ACGAGATCGCGATATTACCGCTGCGCACCTATGGCCCTCCCATGCCAGGGGGCCAATTGC 1200
1201 AGCCCCTCCAGTATTGGCCCTTTTCTTCTGCAGATCTCTATAATTGGAAAACTAACCATC 1260
```

Figure 2 cont.

1261	CCCCTTTCTCGGAGGATCCCCAACGCCTCACGGGGTTGGTGGAGTCCCTTATGTTCTCTC	1320
1321	ACCAGCCTACTTGGGATGATTGTCAACAGCTGCTGCAGACACTCTTCACAACCGAGGAGC	1380
1381	GAGAGAGAATTCTGTTAGAGGCTAGAAAAAATGTTCCTGGGGCCGACGGGCGACCCACGC	1440
1441	AGTTGCAAAATGAGATTGACATGGGATTTCCCTTGACTCGCCCCGGTTGGGACTACAACA	1500
1501	CGGCTGAAGGTAGGGAGAGCTTGAAAATCTATCGCCAGGCTCTGGTGGCGGGTCTCCGGG	1560
1561	GCGCCTCAAGACGGCCCACTAATTTGGCTAAGGTAAGAGAGGTGATGCAGGGACCGAACG	1620
1621	AACCTCCCTCGGTATTTCTTGAGAGGCTCATGGAAGCCTTCAGGCGGTTCACCCCTTTTG	1680
1681	ATCCTACCTCAGAGGCCCAGAAAGCCTCAGTGGCCCTGGCCTTCATTGGGCAGTCGGCTC	1740
1741	TGGATATCAGGAAGAAACTTCAGAGACTGGAAGGGTTACAGGAGGCTGAGTTACGTGATC	1800
1801	TAGTGAGAGAGGCAGAGAAGGTGTATTACAGAAGGGAGACAGAAGAGGAGAAGGAACAGA	1860
1861	GAAAAGAAAAGGAGAGAGAAGAAAGGGAGGAAAGACGTGATAGACGGCAAGAGAAGAATT	1920
1921	TGACTAAGATCTTGGCCGCAGTGGTTGAAGGGAAGAGCAGCAGGGAGAGAGAGAGATT	1980
1981	TTAGGAAAATTAGGTCAGGCCCTAGACAGTCAGGGAACCTGGGCAATAGGACCCCACTCG	2040
2041	ACAAGGACCAGTGTGCGTATTGTAAAGAAAAAGGACACTGGGCAAGGAACTGCCCCAAGA	2100
2101	AGGGAAACAAAGGACCGAAGTCCTAGCTCTAGAAGAAGATAAAGATTAGGGGAGACGGGT	2160
2161	TCGGACCCCCTCCCCGAGCCCAGGGTAACTTTGAAGGTGGAGGGGCAACCAGTTGAGTTC	2220
2221	CTGGTTGATACCGGAGCGGAGCATTCAGTGCTGCTACAACCATTAGGAAAACTAAAAGAA	2280
2281	AAAAAATCCTGGGTGATGGGTGCCACAGGGCAACGGCAGTATCCATGGACTACCCGAAGA	2340
2341	ACCGTTGACTTGGGAGTGGGACGGGTAACCCACTCGTTTCTGGTCATCCCTGAGTGCCCA	2400
2401	GTACCCCTTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAATTTCTTTTGAACAA	2460
2461	GGAAGACCAGAAGTGTCTGTGAATAACAAACCCATCACTGTGTTGACCCTCCAATTAGAT	2520

Figure 2 cont.

2521 GATGAATATCGACTATATTCTCCCCAAGTAAAGCCTGATCAAGATATACAGTCCTGGTTG 2580

2581 GAGCAGTTTCCCCAAGCCTGGGCAGAAACCGCAGGGATGGGTTTGGCAAAGCAAGTTCCC 2640

2641 CCACAGGTTATTCAACTGAAGGCCAGTGCTACACCAGTATCAGTCAGACAGTACCCCTTG 2700

2701 AGTAGAGAGGCTCGAGAAGGAATTTGGCCGCATGTTCAAAGATTAATCCAACAGGGCATC 2760

2761 CTAGTTCCTGTCCAATCCCCTTGGAATACTCCCCTGCTACCGGTTAGGAAGCCTGGGACC 2820

2821 AATGATTATCGACCAGTACAGGACTTGAGAGAGGTCAATAAAAGGGTGCAGGACATACAC 2880

2881 CCAACGGTCCCGAACCCTTATAACCTCTTGAGCGCCCTCCCGCCTGAACGGAACTGGTAC 2940

2941 ACAGTATTGGACTTAAAAGATGCCTTCTTCTGCCTGAGATTACACCCCACTAGCCAACCA 3000

3001 CTTTTTGCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAGCTCACCTGGACC 3060

3061 CGACTGCCCCAAGGGTTCAAGAACTCCCCGACCATCTTTGACGAAGCCCTACACAGGGAC 3120

3121 CTGGCCAACTTCAGGATCCAACACCCTCAGGTGACCCTCCTCCAGTACGTGGATGACCTG 3180

3181 CTTCTGGCGGGAGCCACCAAACAGGACTGCTTAGAAGGTACGAAGGCACTACTGCTGGAA 3240

3241 TTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAGGCCCAGATTTGCAGGAGAGAGGTA 3300

3301 ACATACTTGGGGTACAGTTTGCGGGGCGGGCAGCGATGGCTGACGGAGGCACGGAAGAAA 3360

3361 ACTGTAGTCCAGATACCGGCCCCAACCACAGCCAAACAAGTGAGAGAGTTTTTGGGGACA 3420

3421 GCTGGATTTTGCAGACTGTGGATCCCGGGGTTTGCGACCTTAGCAGCCCCACTCTACCCG 3480

3481 CTAACCAAAGAAAAAGGGGGATTCTCCTGGGCTCCTGAGCACCAGAAGGCATTTGATGCT 3540

3541 ATCAAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACTAAACCCTTT 3600

3601 ACCCTTTATGTGGATGAGCGTAAGGGAGTAGCCCGAGGAGTTTTAACCCAAACCCTAGGA 3660

3661 CCATGGAGGAGACCTGTTGCCTACCTGTCAAAGAAGCTTGATCCTGTAGCCAGTGGTTGG 3720

3721 CCCGTATGTCTGAAGGCTATCGCAGCTGTGGCCATACTGGTCAAGGACGCTGACAAATTG 3780

Figure 2 cont.

3781	ACTTTGGGACAGAATATAACTGTAATAGCCCCCCATGCATTGGAGAACATCGTTCGGCAG	3840
3841	CCCCCAGACCGATGGATGACCAACGCCCGCATGACCCACTATCAAAGCCTGCTTCTCACA	3900
3901	GAGAGGGTCACTTTCGCTCCACCAGCCGCTCTCAACCCTGCCACTCTTCTGCCTGAAGAG	3960
3961	ACTGATGAACCAGTGACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGGTCCGC	4020
4021	AAGGACCTTACAGACATACCGCTGACTGGAGAAGTGCTAACCTGGTTCACTGACGGAAGC	4080
4081	AGCTATGTGGTGGAAGGTAAGAGGATGGCTGGGGCGGCAGTGGTGGACGGGACCCGCACG	4140
4141	ATCTGGGCCAGCAGCCTGCCGGAAGGAACTTCAGCGCAAAAGGCTGAGCTCATGGCCCTC	4200
4201	ACGCAAGCTTTGCGGCTGGCCGAAGGGAAATCCATAAACATTTATACGGACAGCAGGTAT	4260
4261	GCCTTTGCGACTGCACACGTACACGGGGCCATCTATAAACAAAGGGGGTTGCTTACCTCA	4320
4321	GCAGGGAGGGAAATAAAGAACAAAGAGGAAATTCTAAGCCTATTAGAAGCCTTACATTTG	4380
4381	CCAAAAAGGCTAGCTATTATACACTGTCCTGGACATCAGAAAGCCAAAGATCTCATATCT	4440
4441	AGAGGGAACCAGATGGCTGACCGGGTTGCCAAGCAGGCAGCCCAGGCTGTTAACCTTCTG	4500
4501	CCTATAATAGAAACGCCCAAAGCCCCAGAACCCAGACGACAGTACACCCTAGAAGACTGG	4560
4561	CAAGAGATAAAAAAAGATAGACCAGTTCTCTGAGACTCCGGAGGGGACCTGCTATACCTCA	4620
4621	TATGGGAAGGAAATCCTGCCCCACAAAGAAGGGTTAGAATATGTCCAACAGATACATCGT	4680
4681	CTAACCCACCTAGGAACTAAACACCTGCAGCAGTTGGTCAGAACATCCCCTTATCATGTT	4740
4741	CTGAGGCTACCAGGAGTGGCTGACTCGGTGGTCAAACATTGTGTGCCCTGCCAGCTGGTT	4800
4801	AATGCTAATCCTTCCAGAATACCTCCAGGAAAGAGACTAAGGGGAAGCCACCCAGGCGCT	4860
4861	CACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAATACGGAAACAAATATCTATTG	4920
4921	GTTTTTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCCTACTAAGAAAGAGACTTCA	4980
4981	ACCGTGGTGGCTAAGAAAATACTGGAGGAAATTTTTCCAAGATTTGGAATACCTAAGGTA	5040

Figure 2 cont.

5041 ATAGGGTCAGACAATGGTCCAGCTTTCGTTGCCCAGGTAAGTCAGGGACTGGCCAAGATA 5100

5101 TTGGGGATTGATTGGAAACTGCATTGTGCATACAGACCCCAAAGCTCAGGACAGGTAGAG 5160

5161 AGGATGAATAGAACCATTAAAGAGACCCTTACCAAATTGACCACAGAGACTGGCATTAAT 5220

5221 GATTGGATGGCTCTCCTGCCCTTTGTGCTTTTTAGGGTGAGGAACACCCCTGGACAGTTT 5280

5281 GGGCTGACCCCCTATGAATTGCTCTACGGGGGACCCCCCCCGTTGGCAGAAATTGCCTTT 5340

5341 GCACATAGTGCTGATGTGCTGCTTTCCCAGCCTTTGTTCTCTAGGCTCAAGGCGCTCGAG 5400

5401 TGGGTGAGGCAGCGAGCGTGGAAGCAGCTCCGGGAGGCCTACTCAGGAGGAGACTTGCAA 5460

5461 GTTCCACATCGCTTCCAAGTTGGAGATTCAGTCTATGTTAGACGCCACCGTGCAGGAAAC 5520

5521 CTCGAGACTCGGTGGAAGGGACCTTATCTCGTACTTTTGACCACACCAACGGCTGTGAAA 5580

5581 GTCGAAGGAATCCCCACCTGGATCCATGCATCCCACGTTAAGCYGGCGCCACCTCCCGAC 5640

5641 TCGGGGTGGAGAGCCGAAAAGAcTGAGAATCCCCTTAAGCTTCGCCTCCATCGCCTGGTT 5700

5701 CCTTACTCTAACAATAACTCCCCAGGCCAGTAGTAAACGCCTTATAGACAGCTCGAACCC 5760

5761 CCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCCTGATACGGGTGTCACTGTAAA 5820

5821 TAGCACTCGAGGTGTTGCTCCTAGAGGCACCTGGTGGCCTGAACTGCATTTCTGCCTCCG 5880

5881 ATTGATTAACCCCGCTGTTAARAGCACACCTCCCAACCTAGTCCGTAGTTATGGGTTCTA 5940

5941 TTGCTGCCCAGGCACAGAGAAAGAGAAATACTGTGGGGGTTCTGGGGAATCCTTCTGTAG 6000

6001 GAGATGGAGCTGCGTCACCTCCAACGATGGAGACTGGAAATGGCCGATCTCTCTCCAGGA 6060

6061 CCGGGTAAAATTCTCCTTTGTCAATTCCGGCCCGGGCAAGTACAAAATGATGAAACTATA 6120

6121 TAAAGATAAGAGCTGCTCCCCATCAGACTTAGATTATCTAAAGATAAGTTTCACTGAAAG 6180

6181 GAAAACAGGAAAATATTCAAAAGTGGATAAATGGTATGAGCTGGGGAATAGTTTTTTATT 6240

6241 ATATGGCGGGGGAGCAGGGTCCACTTTAACCATTCGCCTTAGGATAGAGACGGGGACAGA 6300

Figure 2 cont.

6301 ACCCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAACAGGGGCCCCCGGCCCTGGA 6360

6361 GCCACCGCATAACTTGCCGGTGCCCCAATTAACCTCGCTGCGGCCTGACATAACACAGCC 6420

6421 GCCTAGCAACAGTACCACTGGATTGATTCCTACCAACACGCCTAGAAACTCCCCAGGTGT 6480

6481 TCCTGTTAAGACAGGACAGAGACTCTTCAGTCTCATCCAGGGAGCTTTCCAAGCCATCAA 6540

6541 CTCCACCGACCCTGATGCCACTTCTTCTTGTTGGCTTTGTCTATCCTCAGGGCCTCCTTA 6600

6601 TTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAGCATAGAAATCAATG 6660

6661 TACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGGAAGGGGACATGCAT 6720

6721 AGGAAAAGCTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTGGTTTATGAGCAGGC 6780

6781 CTCAGAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGGCATGCAATACTGGGTT 6840

6841 AACCCCCTGTGTTTCCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTCATGGTCCA 6900

6901 AATCGTCCCCCGAGTGTACTACCATCCTGAGGAAGTGGTCCTTGATGAATATGACTATCG 6960

6961 GTATAACCGACCAAAAAGAGAACCCGTATCCCTTACCCTAGCTGTAATGCTCGGATTAGG 7020

7021 GACGGCCGTTGGCGTAGGAACAGGGACAGCTGCCCTGATCACAGGACCACAGCAGCTAGA 7080

7081 GAAAGGACTTGGTGAGCTACATGCGGCCATGACAGAAGATCTCCGAGCCTTAAAGGAGTC 7140

7141 TGTTAGCAACCTAGAAGAGTCCCTGACTTCTTTGTCTGAAGTGGTTCTACAGAACCGGAG 7200

7201 GGGATTAGATCTGCTGTTTCTAAGAGAAGGTGGGTTATGTGCAGCCTTAAAAGAAGAATG 7260

7261 TTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCATGAACAAGCTTAGAAAAAA 7320

7321 GTTAGAGAGGCGTCGAAGGGAAAGAGAGGCTGACCAGGGGTGGTTTGAAGGATGGTTCAA 7380

7381 CAGGTCTCCTTGGATGACCACCCTGCTTTCTGCTCTGACGGGGCCCCTAGTAGTCCTGCT 7440

7441 CCTGTTACTTACAGTTGGGCCTTGCTTAATTAATAGGTTTGTTGCCTTTGTTAGAGAACG 7500

7501 AGTGAGTGCAGTCCAGATCATGGTACTTAGGCAACAGTACCAAGGCCTTCTGAGCCAAGG 7560

Figure 2 cont.

```
7561  AGAAACTGACCTCTAGCCTTCCCAGTTCTAAGATTAGAACTATTAACAAGACAAGAAGTG  7620

7621  GGGAATGAAAGGATGAAAATGCAACCTAACCCTCCCAGAACCCAGGAAGTTAATAAAAAG  7680

7681  CTCTAAATGCCCCCGAATTMCAGACCCTGCTGGCTGCCAGTAAATAGGTAGAAGGTCACA  7740

7741  CTTCCTATTGTTCCAGGGCCTGCTATCCTGGCCTAAGTAAGATAACAGGAAATGAGTTGA  7800

7801  CTAATCGCTTATCTGGATTCTGTAAAACTGACTGGCACCATAGAAGAATTGATTACACAT  7860

7861  TGACAGCCCTAGTGACCTATCTCAACTGCAATCTGTCACTCTGCCCAGGAGCCCACGCAG  7920

7921  ATGCGGACCTCCGGAGCTATTTTAAAATGATTGGTCCACGGAGCGCGGGCTCTCGATATT  7980

7981  TTAAAATGATTGGTCCATGGAGCGCGGGCTCTCGATATTTTAAAATGATTGGTTTGTGAC  8040

8041  GCACAGGCTTTGTTGTGAACCCCATAAAAGCTGTCCCGATTCCGCACTCGGGGCCGCAGT  8100

8101  CCTCTACCCCTGCGTGGTGTACGACTGTGGGCCCCAGCGCGCTTGGAATAAAAATCCTCT  8160

8161  TGCTGTTTGCATCAAAAAAAAAAAAAAAAAAAAAA  8196
```

Figure 3


```
   1 GTGGTGTACGACTGTGGGCCCCAGCGCGCTTGGAATAAAAATCCTCTTGCTGTTTGCATC   60

  61 AAGACCGCTTCTCGTGAGTGATTTGGGGTGTCGCCTCTTCCGAGCCCGGACGAGGGGGAT  120

 121 TGTTCTTTTACTGGCCTTTCATTTGGTGCGTTGGCCGGGAAATCCTGCGACCACCCCTTA  180

 181 CACCCGAGAACCGACTTGGAGGTAAAGGGATCCCCTTTGGAACGTGTGTGTGTGTCGGCC  240

 241 GGCGTCTCTGTTCTGAGTGTCTGTTTTCGGTGATGCGCGCTTTCGGTTTGCAGCTGTCCT  300

 301 CTCAGACCGTAAGGACTGGAGGACTGTGATCAGCAGACGTGCTAGGAGGATCACAGGCTG  360

 361 CCACCCTGGGGGACGCCCCGGGAGGTGGGGAGAGCCAGGGACGCCTGGTGGTCTCCTACT  420

 421 GTCGGTCAGAGGACCGAGTTCTGTTGTTGAAGCGAAAGCTTCCCCCTCCGCGGCCGTCCG  480

 481 ACTCTTTTGCCTGCTTGTGGAAGACGCGGACGGGTCGCGTGTGTCTGGATCTGTTGGTTT  540

 541 CTGTCTCGTGTGTCTTTGTCTTGTGCGTCCTTGTCTACAGTTTTAATATGGGACAGACAG  600
                                                    MetGlyGlnThrV

 601 TGACTACCCCCCTTAGTTTGACTCTCGACCATTGGACTGAAGTTAGATCCAGGGCTCATA  660
     alThrThrProLeuSerLeuThrLeuAspHisTrpThrGluValArgSerArgAlaHisA

 661 ATTTGTCAGTTCAGGTTAAGAAGGGACCTTGGCAGACTTTCTGTGCCTCTGAATGGCCAA  720
     snLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysAlaSerGluTrpProT

 721 CATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAATTCTGAAATTATCCTGGCTGTTA  780
     hrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValL

 781 AGGCAATCATTTTTCAGACTGGACCCGGCTCTCATCCTGATCAGGAGCCCTATATCCTTA  840
     ysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIleLeuT

 841 CGTGGCAAGATTTGGCAGAAGATCCTCCGCCATGGGTTAAACCATGGCTAAATAAACCAA  900
     hrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrpLeuAsnLysProA

 901 GAAAGCCAGGTCCCCGAATCCTGGCTCTTGGAGAGAAAAACAAACACTCGGCCGAAAAAG  960
     rgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHisSerAlaGluLysV

 961 TCGAGCCCTCTTCCTCGTATCTACCCCGAGATCGAGGAGCCGCCGACTTGGCCGGAACCC 1020
     alGluProSerSerSerTyrLeuProArgAspArgGlyAlaAlaAspLeuAlaGlyThrP

1021 CAACCTGTTCCCCCACCCCCTTATCCAGCACAGGGTGCTGTGAGGGGACCTCTGCCCCTC 1080
     roThrCysSerProThrProLeuSerSerThrGlyCysCysGluGlyThrSerAlaProP
```

Figure 3 cont.

```
1081 CTGGAGCTCCGGTGGTGGAGGGACCTGCTGCCGGGACTCGGAGCCGGAGAGGCGCCACCC 1140
     roGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAlaThrP

1141 CGGAGCGGACAGACGAGATCGCGATATTACCGCTGCGCACCTATGGCCCTCCCATGCCAG 1200
     roGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGlyProProMetProG

1201 GGGGCCAATTGCAGCCCCTCCAGTATTGGCCCTTTTCTTCTGCAGATCTCTATAATTGGA 1260
     lyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAspLeuTyrAsnTrpL

1261 AAACTAACCATCCCCCTTTCTCGGAGGATCCCCAACGCCTCACGGGGTTGGTGGAGTCCC 1320
     ysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGlyLeuValGluSerL

1321 TTATGTTCTCTCACCAGCCTACTTGGGATGATTGTCAACAGCTGCTGCAGACACTCTTCA 1380
     euMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheT

1381 CAACCGAGGAGCGAGAGAGAATTCTGTTAGAGGCTAGAAAAAATGTTCCTGGGGCCGACG 1440
     hrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAlaAspG

1441 GGCGACCCACGCAGTTGCAAAATGAGATTGACATGGGATTTCCCTTGACTCGCCCCGGTT 1500
     lyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeuThrArgProGlyT

1501 GGGACTACAACACGGCTGAAGGTAGGGAGAGCTTGAAAATCTATCGCCAGGCTCTGGTGG 1560
     rpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArgGlnAlaLeuValA

1561 CGGGTCTCCGGGGCGCCTCAAGACGGCCCACTAATTTGGCTAAGGTAAGAGAGGTGATGC 1620
     laGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysValArgGluValMetG

1621 AGGGACCGAACGAACCTCCCTCGGTATTTCTTGAGAGGCTCATGGAAGCCTTCAGGCGGT 1680
     lnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGluAlaPheArgArgP

1681 TCACCCCTTTTGATCCTACCTCAGAGGCCCAGAAAGCCTCAGTGGCCCTGGCCTTCATTG 1740
     heThrProPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPheIleG

1741 GGCAGTCGGCTCTGGATATCAGGAAGAAACTTCAGAGACTGGAAGGGTTACAGGAGGCTG 1800
     lyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGlyLeuGlnGluAlaG

1801 AGTTACGTGATCTAGTGAGAGAGGCAGAGAAGGTGTATTACAGAAGGGAGACAGAAGAGG 1860
     luLeuArgAspLeuValArgGluAlaGluLysValTyrTyrArgArgGluThrGluGluG

1861 AGAAGGAACAGAGAAAAGAAAAGGAGAGAGAAGAAAGGGAGGAAAGACGTGATAGACGGC 1920
     luLysGluGlnArgLysGluLysGluArgGluGluArgGluGluArgArgAspArgArgG

1921 AAGAGAAGAATTTGACTAAGATCTTGGCCGCAGTGGTTGAAGGGAAGAGCAGCAGGGAGA 1980
     lnGluLysAsnLeuThrLysIleLeuAlaAlaValValGluGlyLysSerSerArgGluA

1981 GAGAGAGAGATTTTAGGAAAATTAGGTCAGGCCCTAGACAGTCAGGGAACCTGGGCAATA 2040
     rgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGlyAsnA
```

Figure 3 cont

```
2041 GGACCCCACTCGACAAGGACCAGTGTGCGTATTGTAAAGAAAAAGGACACTGGGCAAGGA 2100
     rgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGlyHisTrpAlaArgA

2101 ACTGCCCCAAGAAGGGAAACAAAGGACCGAAGgTCCTAGCTCTAGAAGAAGATAAAGATT 2160
     snCysProLysLysGlyAsnLysGlyProLysValLeuAlaLeuGluGluAspLysAspE

2161 AGGGGAGACGGGgTTCGGACCCCCTCCCCGAGCCCAGGGTAACTTTGAAGGTGGAGGGGC 2220
     ndGlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysValGluGlyG

2221 AACCAGTTGAGTTCCTGGTTGATACCGGAGCGGAGCATTCAGTGCTGCTACAACCATTAG 2280
     lnProValGluPheLeuValAspThrGlyAlaGluHisSerValLeuLeuGlnProLeuG

2281 GAAAACTAAAAGAAAAAAAATCCTGGGTGATGGGTGCCACAGGGCAACGGCAGTATCCAT 2340
     lyLysLeuLysGluLysLysSerTrpValMetGlyAlaThrGlyGlnArgGlnTyrProT

2341 GGACTACCCGAAGAACCGTTGACTTGGGAGTGGGACGGGTAACCCACTCGTTTCTGGTCA 2400
     rpThrThrArgArgThrValAspLeuGlyValGlyArgValThrHisSerPheLeuValI

2401 TCCCTGAGTGCCcAGTACCCCTTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAA 2460
     leProGluCysProValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnI

2461 TTTCTTTTGAACAAGGAAGACCAGAAGTGTCTGTGAATAACAAACCCATCACTGTGTTGA 2520
     leSerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleThrValLeuT

2521 CCCTCCAATTAGATGATGAATATCGACTATATTCTCCCCAAGTAAAGCCTGATCAAGATA 2580
     hrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnValLysProAspGlnAspI

2581 TACAGTCCTGGTTGGAGCAGTTTCCCCAAGCCTGGGCAGAAACCGCAGGGATGGGTTTGG 2640
     leGlnSerTrpLeuGluGlnPheProGlnAlaTrpAlaGluThrAlaGlyMetGlyLeuA

2641 CAAAGCAAGTTCCCCCACAGGTTATTCAACTGAAGGCCAGTGCTACACCAGTATCAGTCA 2700
     laLysGlnValProProGlnValIleGlnLeuLysAlaSerAlaThrProValSerValA

2701 GACAGTACCCCTTGAGTAGAGAGGCTCGAGAAGGAATTTGGCCGCATGTTCAAAGATTAA 2760
     rgGlnTyrProLeuSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuI

2761 TCCAACAGGGCATCCTAGTTCCTGTCCAATCCCCTTGGAATACTCCCCTGCTACCGGTTA 2820
     leGlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLeuProValA

2821 GGAAGCCTGGGACCAATGATTATCGACCAGTACAGGACTTGAGAGAGGTCAATAAAAGGG 2880
     rgLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsnLysArgV

2881 TGCAGGACATACACCCAACGGTCCCGAACCCTTATAACCTCTTGAGCGCCCTCCCGCCTG 2940
     alGlnAspIleHisProThrValProAsnProTyrAsnLeuLeuSerAlaLeuProProG

2941 AACGGAACTGGTACACAGTATTGGACTTAAAAGATGCCTTCTTCTGCCTGAGATTACACC 3000
     luArgAsnTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeuHisP
```

Figure 3 cont.

```
3001 CCACTAGCCAACCACTTTTTGCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGC 3060
     roThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyG

3061 AGCTCACCTGGACCCGACTGCCCCAAGGGTTCAAGAACTCCCCGACCATCTTTGACGAAG 3120
     lnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePheAspGluA

3121 CCCTACACAGGGACCTGGCCAACTTCAGGATCCAACACCCTCAGGTGACCCTCCTCCAGT 3180
     laLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnValThrLeuLeuGlnT

3181 ACGTGGATGACCTGCTTCTGGCGGGAGCCACCAAACAGGACTGCTTAGAAGGTACGAAGG 3240
     yrValAspAspLeuLeuLeuAlaGlyAlaThrLysGlnAspCysLeuGluGlyThrLysA

3241 CACTACTGCTGGAATTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAGGCCCAGATTT 3300
     laLeuLeuLeuGluLeuSerAspLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleC

3301 GCAGGAGAGAGGTAACATACTTGGGGTACAGTTTGCGGGGCGGGCAGCGATGGCTGACGG 3360
     ysArgArgGluValThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrG

3361 AGGCACGGAAGAAAACTGTAGTCCAGATACCGGCCCCAACCACAGCCAAACAAGTGAGAG 3420
     luAlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGlnValArgG

3421 AGTTTTTGGGGACAGCTGGATTTTGCAGACTGTGGATCCCGGGGTTTGCGACCTTAGCAG 3480
     luPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaThrLeuAlaA

3481 CCCCACTCTACCCGCTAACCAAAGAAAAAGGGGGATTCTCCTGGGCTCCTGAGCACCAGA 3540
     laProLeuTyrProLeuThrLysGluLysGlyGlyPheSerTrpAlaProGluHisGlnL

3541 AGGCATTTGATGCTATCAAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACG 3600
     ysAlaPheAspAlaIleLysLysAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspV

3601 TAACTAAACCCTTTACCCTTTATGTGGATGAGCGTAAGGGAGTAGCCCGAGGAGTTTTAA 3660
     alThrLysProPheThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuT

3661 CCCAAACCCTAGGACCATGGAGGAGACCTGTTGCCTACCTGTCAAAGAAGCTTGATCCTG 3720
     hrGlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLeuAspProV

3721 TAGCCAGTGGTTGGCCCGTATGTCTGAAGGCTATCGCAGCTGTGGCCATACTGGTCAAGG 3780
     alAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaValAlaIleLeuValLysA

3781 ACGCTGACAAATTGACTTTGGGACAGAATATAACTGTAATAGCCCCCCATGCATTGGAGA 3840
     spAlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleAlaProHisAlaLeuGluA

3841 ACATCGTTCGGCAGCCCCCAGACCGATGGATGACCAACGCCCGCATGACCCACTATCAAA 3900
     snIleValArgGlnProProAspArgTrpMetThrAsnAlaArgMetThrHisTyrGlnS

3901 GCCTGCTTCTCACAGAGAGGGTCACTTTCGCTCCACCAGCCGCTCTCAACCCTGCCACTC 3960
     erLeuLeuLeuThrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrL
```

Figure 3 cont

```
3961 TTCTGCCTGAAGAGACTGATGAACCAGTGACTCATGATTGCCATCAACTATTGATTGAGG 4020
     euLeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLeuIleGluG

4021 AGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTGGAGAAGTGCTAACCTGGT 4080
     luThrGlyValArgLysAspLeuThrAspIleProLeuThrGlyGluValLeuThrTrpP

4081 TCACTGACGGAAGCAGCTATGTGGTGGAAGGTAAGAGGATGGCTGGGGCGGCAGTGGTGG 4140
     heThrAspGlySerSerTyrValValGluGlyLysArgMetAlaGlyAlaAlaValValA

4141 ACGGGACCCGCACGATCTGGGCCAGCAGCCTGCCGGAAGGAACTTCAGCGCAAAAGGCTG 4200
     spGlyThrArgThrIleTrpAlaSerSerLeuProGluGlyThrSerAlaGlnLysAlaG

4201 AGCTCATGGCCCTCACGCAAGCTTTGCGGCTGGCCGAAGGGAAATCCATAAACATTTATA 4260
     luLeuMetAlaLeuThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrT

4261 CGGACAGCAGGTATGCCTTTGCGACTGCACACGTACACGGGGCCATCTATAAACAAAGGG 4320
     hrAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLysGlnArgG

4321 GGTTGCTTACCTCAGCAGGGAGGGAAATAAAGAACAAAGAGGAAATTCTAAGCCTATTAG 4380
     lyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluGluIleLeuSerLeuLeuG

4381 AAGCCTTACATTTGCCAAAAAGGCTAGCTATTATACACTGTCCTGGACATCAGAAAGCCA 4440
     luAlaLeuHisLeuProLysArgLeuAlaIleIleHisCysProGlyHisGlnLysAlaL

4441 AAGATCTCATATCTAGAGGGAACCAGATGGCTGACCGGGTTGCCAAGCAGGCAGCCCAGG 4500
     ysAspLeuIleSerArgGlyAsnGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnA

4501 CTGTTAACCTTCTGCCTATAATAGAAACGCCCAAAGCCCCAGAACCCAGACGACAGTACA 4560
     laValAsnLeuLeuProIleIleGluThrProLysAlaProGluProArgArgGlnTyrT

4561 CCCTAGAAGACTGGCAAGAGATAAAAAAAGATAGACCAGTTCTCTGAGACTCCGGAGGGGA 4620
     hrLeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrProGluGlyT

4621 CCTGCTATACCTCATATGGGAAGGAAATCCTGCCCCACAAAGAAGGGTTAGAATATGTCC 4680
     hrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysGluGlyLeuGluTyrValG

4681 AACAGATACATCGTCTAACCCACCTAGGAACTAAACACCTGCAGCAGTTGGTCAGAACAT 4740
     lnGlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuGlnGlnLeuValArgThrS

4741 CCCCTTATCATGTTCTGAGGCTACCAGGAGTGGCTGACTCGGTGGTCAAACATTGTGTGC 4800
     erProTyrHisValLeuArgLeuProGlyValAlaAspSerValValLysHisCysValP

4801 CCTGCCAGCTGGTTAATGCTAATCCTTCCAGAATACCTCCAGGAAAGAGACTAAGGGGAA 4860
     roCysGlnLeuValAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlyS

4861 GCCACCCAGGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAATACGGAA 4920
     erHisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLysTyrGlyA
```

Figure 3 cont.

```
4921  ACAAATATCTATTGGTTTTTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCCTACTA  4980
      snLysTyrLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaTyrProThrL

4981  AGAAAGAGACTTCAACCGTGGTGGCTAAGAAAATACTGGAGGAAATTTTTCCAAGATTTG  5040
      ysLysGluThrSerThrValValAlaLysLysIleLeuGluGluIlePheProArgPheG

5041  GAATACCTAAGGTAATAGGGTCAGACAATGGTCCAGCTTTCGTTGCCCAGGTAAGTCAGG  5100
      lyIleProLysValIleGlySerAspAsnGlyProAlaPheValAlaGlnValSerGlnG

5101  GACTGGCCAAGATATTGGGGATTGATTGGAAACTGCATTGTGCATACAGACCCCAAAGCT  5160
      lyLeuAlaLysIleLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerS

5161  CAGGACAGGTAGAGAGGATGAATAGAACCATTAAAGAGACCCTTACCAAATTGACCACAG  5220
      erGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThrThrG

5221  AGACTGGCATTAATGATTGGATGGCTCTCCTGCCCTTTGTGCTTTTTAGGGTGAGGAACA  5280
      luThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValLeuPheArgValArgAsnT

5281  CCCCTGGACAGTTTGGGCTGACCCCCTATGAATTGCTCTACGGGGGACCCCCCCCGTTGG  5340
      hrProGlyGlnPheGlyLeuThrProTyrGluLeuLeuTyrGlyGlyProProProLeuA

5341  CAGAAATTGCCTTTGCACATAGTGCTGATGTGCTGCTTTCCCAGCCTTTGTTCTCTAGGC  5400
      laGluIleAlaPheAlaHisSerAlaAspValLeuLeuSerGlnProLeuPheSerArgL

5401  TCAAGGCGCTCGAGTGGGTGAGGCAGCGAGCGTGGAAGCAGCTCCGGGAGGCCTACTCAG  5460
      euLysAlaLeuGluTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerG

5461  GAGGAGACTTGCAAGTTCCACATCGCTTCCAAGTTGGAGATTCAGTCTATGTTAGACGCC  5520
      lyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrValArgArgH

5521  ACCGTGCAGGAAACCTCGAGACTCGGTGGAAGGGACCTTATCTCGTACTTTTGACCACAC  5580
      isArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrLeuValLeuLeuThrThrP

5581  CAACGGCTGTGAAAGTCGAAGGAATCCCCACCTGGATCCATGCATCCCACGTTAAGCCGG  5640
      roThrAlaValLysValGluGlyIleProThrTrpIleHisAlaSerHisValLysProA
                                                    MetHisProThrLeuSerArg

5641  CGCCACCTCCCGACTCGGGGTGGAGAGCCGAAAAGAcTGAGAATCCCCTTAAGCTTCGCC  5700
      laProProProAspSerGlyTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgL
      ArgHisLeuProThrArgGlyGlyGluProLysArgLeuArgIleProLeuSerPheAla

5701  TCCATCGCCTGGTTCCTTACTCTAACAATAACTCCCCAGGCCAGTAGTAAACGCCTTATA  5760
      euHisArgLeuValProTyrSerAsnAsnAsnSerProGlyGlnEnd
      SerIleAlaTrpPheLeuThrLeuThrIleThrProGlnAlaSerSerLysArgLeuIle

5761  GACAGCTCGAACCCCCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCCTGATACG  5820
      AspSerSerAsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspThr
```

Figure 3 cont

```
5821  GGTGTCACTGTAAATAGCACTCGAGGTGTTGCTCCTAGAGGCACCTGGTGGCCTGAACTG  5880
      GlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpTrpProGluLeu

5881  CATTTCTGCCTCCGATTGATTAACCCCGCTGTTAAAAGCACACCTCCCAACCTAGTCCGT  5940
      HisPheCysLeuArgLeuIleAsnProAlaValLysSerThrProProAsnLeuValArg

5941  AGTTATGGGTTCTATTGCTGCCCAGGCACAGAGAAAGAGAAATACTGTGGGGGTTCTGGG  6000
      SerTyrGlyPheTyrCysCysProGlyThrGluLysGluLysTyrCysGlyGlySerGly

6001  GAATCCTTCTGTAGGAGATGGAGCTGCGTCACCTCCAACGATGGAGACTGGAAATGGCCG  6060
      GluSerPheCysArgArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro

6061  ATCTCTCTCCAGGACCGGGTAAAATTCTCCTTTGTCAATTCCGGCCCGGGCAAGTACAAA  6120
      IleSerLeuGlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLys

6121  ATGATGAAACTATATAAAGATAAGAGCTGCTCCCCATCAGACTTAGATTATCTAAAGATA  6180
      MetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspTyrLeuLysIle

6181  AGTTTCACTGAAAGGAAAACAGGAAAATATTCAAAAGTGGATAAATGGTATGAGCTGGGG  6240
      SerPheThrGluArgLysThrGlyLysTyrSerLysValAspLysTrpTyrGluLeuGly

6241  AATAGTTTTTTATTATATGGCGGGGGAGCAGGGTCCACTTTAACCATTCGCCTTAGGATA  6300
      AsnSerPheLeuLeuTyrGlyGlyGlyAlaGlySerThrLeuThrIleArgLeuArgIle

6301  GAGACGGGGACAGAACCCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAACAGGGG  6360
      GluThrGlyThrGluProProValAlaMetGlyProAspLysValLeuAlaGluGlnGly

6361  CCCCCGGCCCTGGAGCCACCGCATAACTTGCCGGTGCCCCAATTAACCTCGCTGCGGCCT  6420
      ProProAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuArgPro

6421  GACATAACACAGCCGCCTAGCAACAGTACCACTGGATTGATTCCTACCAACACGCCTAGA  6480
      AspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProThrAsnThrProArg

6481  AACTCCCCAGGTGTTCCTGTTAAGACAGGACAGAGACTCTTCAGTCTCATCCAGGGAGCT  6540
      AsnSerProGlyValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAla

6541  TTCCAAGCCATCAACTCCACCGACCCTGATGCCACTTCTTCTTGTTGGCTTTGTCTATCC  6600
      PheGlnAlaIleAsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSer

6601  TCAGGGCCTCCTTATTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAG  6660
      SerGlyProProTyrTyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGlu

6661  CATAGAAATCAATGTACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGG  6720
      HisArgAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGly

6721  AAGGGGACATGCATAGGAAAAGCTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTG  6780
      LysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuCysTyrSerThrVal
```

Figure 3 cont. 18/22

```
6781  GTTTATGAGCAGGCCTCAGAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGGCA  6840
      ValTyrGluGlnAlaSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAla

6841  TGCAATACTGGGTTAACCCCCTGTGTTTCCACCTCAGTCTTCAACCAATCCAAAGATTTC  6900
      CysAsnThrGlyLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPhe

6901  TGTGTCATGGTCCAAATCGTCCCCCGAGTGTACTACCATCCTGAGGAAGTGGTCCTTGAT  6960
      CysValMetValGlnIleValProArgValTyrTyrHisProGluGluValValLeuAsp

6961  GAATATGACTATCGGTATAACCGACCAAAAAGAGAACCCGTATCCCTTACCCTAGCTGTA  7020
      GluTyrAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaVal

7021  ATGCTCGGATTAGGGACGGCCGTTGGCGTAGGAACAGGGACAGCTGCCCTGATCACAGGA  7080
      MetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaAlaLeuIleThrGly

7081  CCACAGCAGCTAGAGAAAGGACTTGGTGAGCTACATGCGGCCATGACAGAAGATCTCCGA  7140
      ProGlnGlnLeuGluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArg

7141  GCCTTAAAGGAGTCTGTTAGCAACCTAGAAGAGTCCCTGACTTCTTTGTCTGAAGTGGTT  7200
      AlaLeuLysGluSerValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValVal

7201  CTACAGAACCGGAGGGGATTAGATCTGCTGTTTCTAAGAGAAGGTGGGTTATGTGCAGCC  7260
      LeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAla

7261  TTAAAAGAAGAATGTTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCATGAAC  7320
      LeuLysGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetAsn

7321  AAGCTTAGAAAAAAGTTAGAGAGGCGTCGAAGGGAAAGAGAGGCTGACCAGGGGTGGTTT  7380
      LysLeuArgLysLysLeuGluArgArgArgArgGluArgGluAlaAspGlnGlyTrpPhe

7381  GAAGGATGGTTCAACAGGTCTCCTTGGATGACCACCCTGCTTTCTGCTCTGACGGGGCCC  7440
      GluGlyTrpPheAsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyPro

7441  CTAGTAGTCCTGCTCCTGTTACTTACAGTTGGGCCTTGCTTAATTAATAGGTTTGTTGCC  7500
      LeuValValLeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAla
      -

7501  TTTGTTAGAGAACGAGTGAGTGCAGTCCAGATCATGGTACTTAGGCAACAGTACCAAGGC  7560
      PheValArgGluArgValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGly

7561  CTTCTGAGCCAAGGAGAAACTGACCTCTAGCCTTCCCAGTTCTAAGATTAGAACTATTAA  7620
      LeuLeuSerGlnGlyGluThrAspLeuEnd

7621  CAAGACAAGAAGTGGGGAATGAAAGGATGAAAATGCAACCTAACCCTCCCAGAACCCAGG  7680

7681  AAGTTAATAAAAAGCTCTAAATGCCCCCGAATTACAGACCCTGCTGGCTGCCAGTAAATA  7740
```

Figure 3 cont.

7741 GGTAGAAGGTCACACTTCCTATTGTTCCAGGGCCTGCTATCCTGGCCTAAGTAAGATAAC 7800

7801 AGGAAATGAGTTGACTAATCGCTTATCTGGATTCTGTAAAACTGACTGGCACCATAGAAG 7860

7861 AATTGATTACACATTGACAGCCCTAGTGACCTATCTCAACTGCAATCTGTCACTCTGCCC 7920

7921 AGGAGCCCACGCAGATGCGGACCTCCGGAGCTATTTTAAAATGATTGGTCCACGGAGCGC 7980

7981 GGGCTCTCGATATTTTAAAATGATTGGTCCATGGAGCGCGGGCTCTCGATATTTTAAAAT 8040

8041 GATTGGTTTGTGACGCACAGGCTTTGTTGTGAACCCCATAAAAGCTGTCCCGATTCCGCA 8100

8101 CTCGGGGCCGCAGTCCTCTACCCCTGCGTGGTGTACGACTGTGGGCCCCAGCGCGCTTGG 8160

8161 AATAAAAATCCTCTTGCTGTTTGCATCAAAAAAAAAAAAAAAAAAAAA 8209

Figure 4.

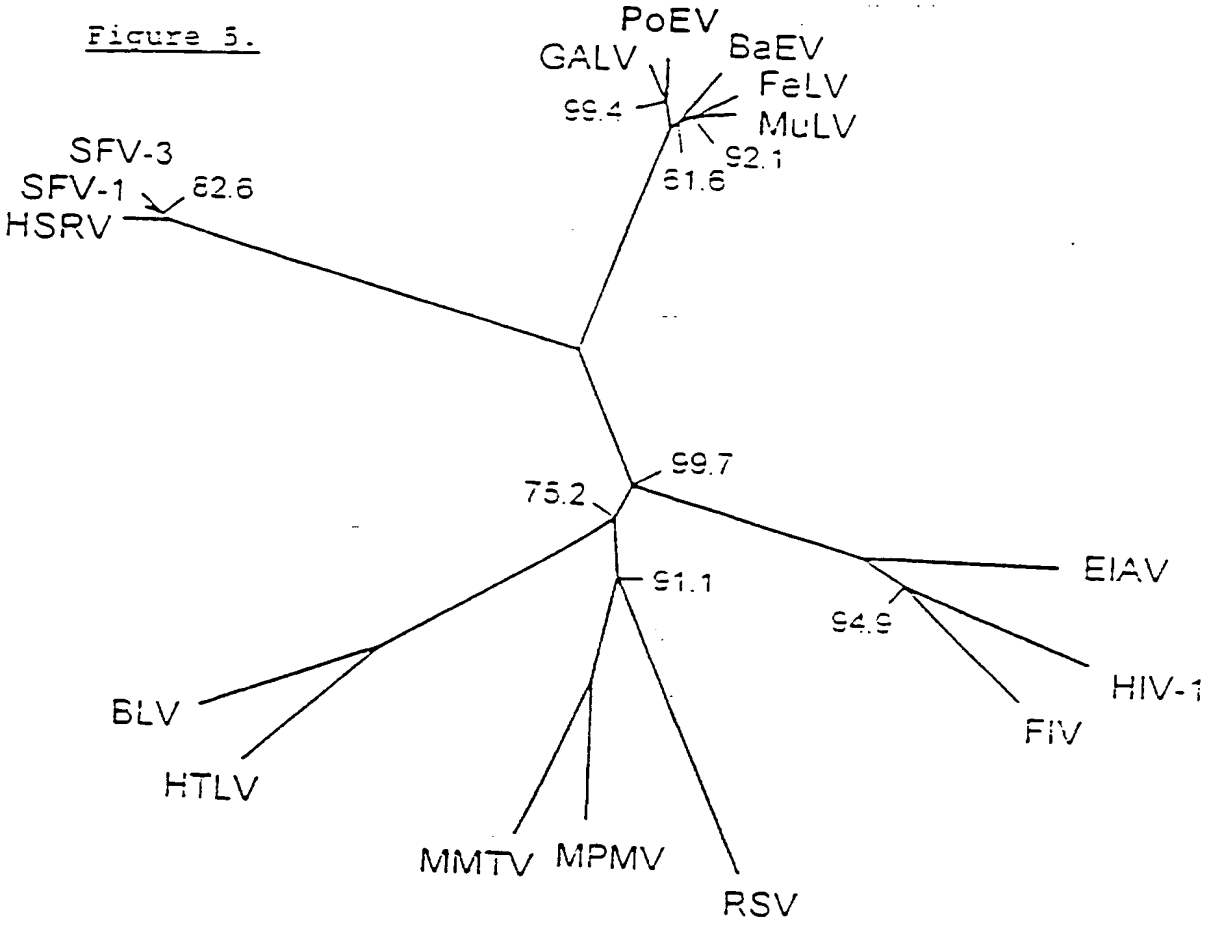
The same nucleotide sequence as represented by bases 5260 to 8210 in Figure 3 is also representative for this Figure, with the following changes:

Position	Change
5273	G-T
5341	C-T
5351	C-T
5353	T-C
5356	C-T
5426	G-A
5464	Insertion AGA
5607	C-T
5638	C-T
5792	T-C
6191	Insertion AA
6253	T-A
6255	Insertion A
6900	C-G

Such nucleotide changes result in the following amino acid changes in the ENV polypeptide.

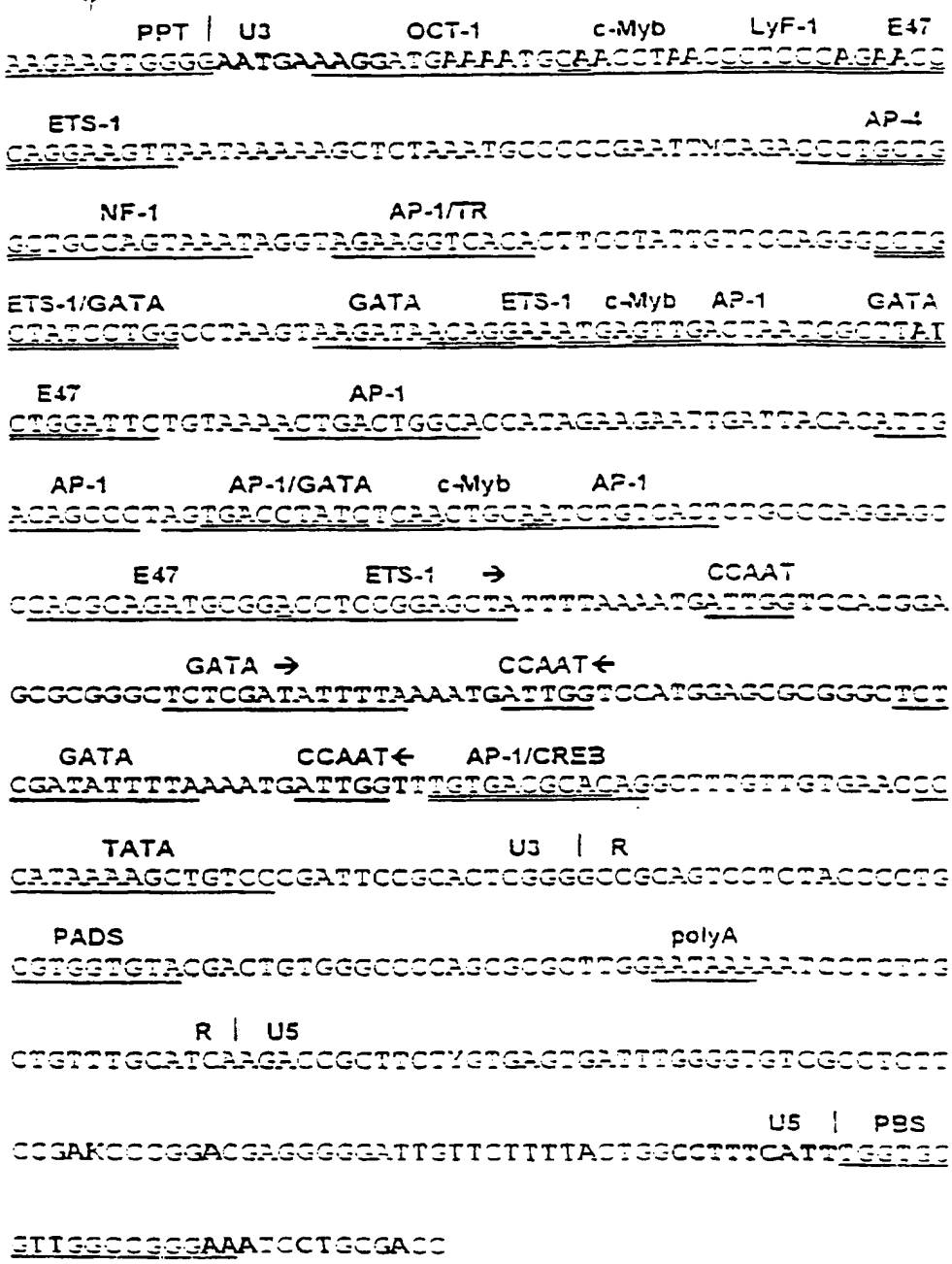
Position	Change
7	R-W
192	R-K
193	Deletion
194	Deletion
197	Y-Q
198	S-E
199	K-N
200	V-I
201	D-Q
204	Y-I
205	E-N
206	Insertions: G,M,S
206	L-W
208	N-I
209	S-V
211	L-Y
212	L-K
427	F-L

Figure 5.



MuLV murine leukaemia virus
FeLV feline leukaemia virus
GaLV gibbon ape leukaemia virus
SVV-1 simian foamy virus 1
SFV-3 simian foamy virus 3
HSRV human foamy virus
BLV Bovine leukaemia virus
HTLV human T-cell leukaemia virus
MMTV murine mammary tumour virus
MPMV Mason Pfizer monkey virus
RSV Rous sarcoma virus
FIV feline immunodeficiency virus
HIV human immunodeficiency virus
EIAV equine infectious anaemia virus

Figure 6



INTERNATIONAL SEARCH REPORT

Inter. nal Application No
PCT/GB 97/01087

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/48 C07K14/15 A61K39/21 C07K16/10 C12Q1/70
A01K67/027

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A61K C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	WO 97 21836 A (THE GENERAL HOSPITAL CORPORATION) 19 June 1997 see the whole document, especially SEQ ID NOs:2 and 3 and claims 1-38 ---	1-31
A	NATURE MEDICINE, vol. 1, no. 11, November 1995, page 1100 XP002037073 J.P. STOYE AND J.M. COFFIN: "The dangers of xenotransplantation" see the whole document ---	1-31
P,X	NATURE MEDICINE, vol. 3, no. 3, March 1997, pages 282-286, XP002037074 C. PATIENCE ET AL.: "Infection of human cells by an endogenous retrovirus of pigs" see the whole document -----	1-24,31

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
7 August 1997	14. 08. 97

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	Cupido, M

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

ternational application No.
PCT/GB 97/01087

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 29
because they relate to subject matter not required to be searched by this Authority, namely:
Although this claim, insofar in vivo uses are concerned, is directed to a method of treatment and a diagnostic method practised on the human/ animal body the search has been carried out based on the alleged effects of the compounds.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

information on patent family members

Inter nal Application No
PCT/GB 97/01087

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9721836 A	19-06-97	NONE	